MARINTEK

Norsk Marinteknisk Forskningsinstitutt AS
Norwegian Marine Technology Research Institute

Address:
POB 4125 Valentinlyst
NO-7450 Trondheim
NORWAY

Telephone: +47 464 15 000
Telefax:+47 73595778

marintek@marintek.sintef.no
www.marintek.sintef.no
Enterprise /VAT No:
NO 937 357 370 MVA

Memo D0 - SIMVAL

State of the art document on Validation of simulation models for ship manoeuvring

PERSON RESPONSIBLE / AUTHOR
Tor Einar Berg

DISTRIBUTION	FOR YOUR ATTENTION	COMMENTS ARE INVITED	FOR YOUR INFORMATION	AS AGREED
Project partners				X

PROJECT NO / FILE CODE	DATE	CLASSIFICATION
302000111	2015-05-20	Unrestricted

Abstract

This memo contains a review of previous work on the topic validation of simulation models for ship manoeuvring. In addition, some papers describing verification, validation and accreditation in general are reviewed. Even though the concepts of verification and validation often are used in discussions of simulator fidelity no one in the maritime simulation community has come up with a good description of how to validate shiphandling simulation models.

A draft version of the memo was circulated for comments by project partners. Comments received by 20th June 2014 were included in the August 2014 version of the document. This final version has been prepared to include the outcomes from the 27th ITTC held in Copenhagen August 2014 and the SIMMAN 2014 workshop in December 2014.

MARINTEK

Table of Contents

1 Introduction

Previously the Mariner and Series-60 classes of vessels have been used as reference cases for studies of ship manoeuvring performance. Based on theoretical models and captive model tests simulation models were developed. Validation data based on sea trials with these vessel classes are not easily available. Some early studies have been reviewed by the Manoeuvring Committee of the International Towing Tank Conference (ITTC).

In the late 1970-ties Exxon sponsored a major sea trial program for their VLCC "Esso Osaka". Results from these tests have been published by Lincoln Crane et al (1979). Ship lines have been made available for research organisations which have built scaled models and developed numerical simulation models for this VLCC. "Esso Osaka" has been used as a case study vessel by ITTC's Manoeuvring Committee and findings from simulation of sea trials have been published in ITTC reports (16^{th} – 18^{th} ITTC Manoeuvring Committee). It has been found that there are some uncertainties in the environmental conditions during the sea trials with Esso Osaka making it difficult to use these trials for validation of simulation models. As a consequence the ITTC Manoeuvring Committee decided to take another approach for validation studies. A number of concept vessels were selected for comparison of manoeuvring performance based on model test results and/or numerical models. Simulation models based on empirical methods, CFD codes and captive models tests were used to calculate IMO's standard manoeuvres for deep water in calm conditions. A large number of research companies participated in a blind test study and the results were presented at the SIMMAN 2008 workshop in Copenhagen in April 2008. The results showed large variations between research organisations and between methods used for simulation of the standard manoeuvres (Stern and Agdrup, 2008). The organisations behind SIMMAN 2008 have for some years planned a follow-up workshop. This will take place in December 2014 and for this workshop the R&D companies are asked to add shallow water performance to their simulations. Flanders Hydraulics Research is one of the organisations that will provide captive model test data from shallow water PMM tests.

As none of the SIMMAN case vessels has been built, their validation study is limited to comparing simulated results to results obtained from free-sailing models. The scaling problem going from model scale to the real ship is thus not a part of the SIMMAN study.

2 Some general comments on verification and validation.

To start off the project it was decided to run validation workshops. The first one was done using video conferencing, the second one was organized by Flanders Hydraulics Research early October 2013. At the first workshop the background documents were the ITTC Recommended Procedures and Guidelines relevant for studies of ship manoeuvrability and an OMAE paper by Berg and Ringen (Berg and Ringen, 2011). In this paper they state that evaluation of simulation systems used for engineering studies to a certain degree is parallel to the evaluation needed for training simulators. For both applications, the major items to be investigated are:

- Simulation fidelity – level of realism that the simulator is presenting to the user.
- Simulation verification – the process of ascertain that the mathematical model is operating as intended.
- Simulation validation – the process of ensuring that conclusions reached from simulation are similar to those reached in the real world system being modelled.

At the second workshop Vantorre gave an overview of validation work reported by different ITTC Committees. In the 2011 Terms of Reference for the Quality Systems Group it is stated that they should:

- Include a definition of the terms Verification and Validation in the ITTC Symbols and Terminology List
- This should be done within first three months as a basis for the work of other committees.

This activity has not been completed yet, it is expected that a text proposal will be presented at the 2014 ITTC meeting. At the same meeting Berg presented some of the content in a paper by Bair on a more generic validation process (Bair, 2012). Here she describes "The Validation Process Maturity Model (VPMM)" developed by Harmon and Youngblood (2005). It addresses the fundamental truth that validation assessments differ in their levels of objectivity, repeatability, timeliness, completeness, and accuracy. Its function is to provide a means to assess the validation process and the completeness of a validation assessment as proxies for the quality of the information provided.

Starting in the 1990-ties, government agencies and professional societies increased their focus on verification, validation and accreditation (VV&A) significantly. A paper by Page (Page, 2004) stated that there are a variety of formal definitions of VV&A terms. In his paper Page made use of the following definitions:

- Verification
  - Did I build the thing right?
  - Have the model and simulation been built so that they fully satisfy the developer's intent (as indicated in specifications)?
  - Verification has two aspects: design (all specifications and nothing else are included in the model or simulation design) and implementation (all specifications and nothing else are included in the model or simulation as built).
- Validation
  - Did I build the right thing?
  - Will the model or simulation be able to adequately support its intended use?
  - Is its fidelity appropriate for that?
  - Validation has two aspects: conceptual validation (when the anticipated fidelity of the model or simulation conceptual model is assessed) and results validation (when results from the implemented model or simulation are compared with an appropriate referent to demonstrate that the model or simulation can in fact support the intended use).
- Accreditation

- Accreditation is a management decision that may include schedule and other considerations as well as technical V&V information.
- Authority for the accreditation decision is normally vested in those responsible for consequences from the use of results from the model or simulation.
- Often the accreditation decision is based on whether the model or simulation can support a subset of its requirements called “acceptability criteria”.

Figure 1 illustrates real-world and simulation-world relationships in developing system theories and simulation models with verification and validation.

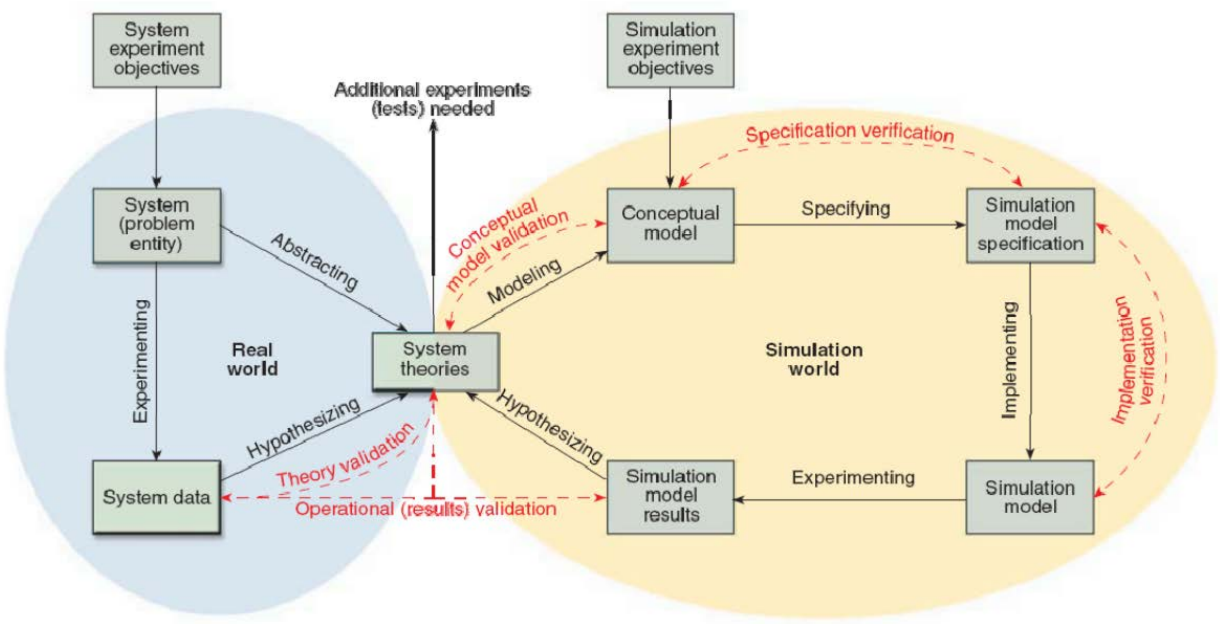

Figure 1 – Sargent Circle of Validation: diagram developed and copyrighted by Dr. R.G. Sargent, Syracuse University, January 2001. Used with Permission.

Figure 2 shows relationships for requirements, acceptability criteria, verification, validation and other information used in an accreditation decision. Pace also points to the growing reliance on modelling and simulation results, both to guide system design and to develop operational philosophies. This makes it ever more important to make simulated results acceptable under all circumstances and to know clearly where the limits are on their applicability.

In a conference paper from 2011 (Sargent, 2011) it is stated that it is often too costly and time consuming to determine that a model is absolutely valid over the complete domain of its intended applicability. Instead, tests and evaluations are conducted until sufficient confidence is obtained that a model can be considered valid for its intended application. If a test determines that a model does not have sufficient accuracy for any one of the sets of experimental conditions, then the model is invalid. However, determining that a model has sufficient accuracy for numerous experimental conditions does not guarantee that a model is valid everywhere in its applicable domain. Figure 3 shows the relationships between model confidence and (a) cost (a similar relationship holds for the amount of time) of performing model validation and (b) the value of the model to a user. The cost of model validation is usually quite significant, especially when extremely high model confidence is required. Sargent recommends that, as a minimum, the following eight steps should be performed in model validation:

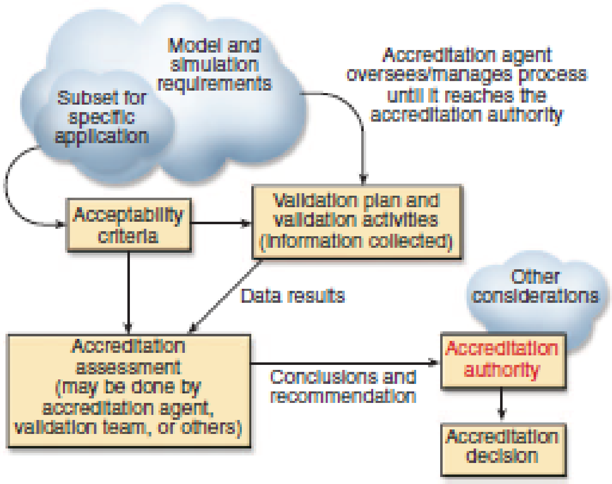

Figure 2. Overview of the accreditation process (from Pace, 2004).

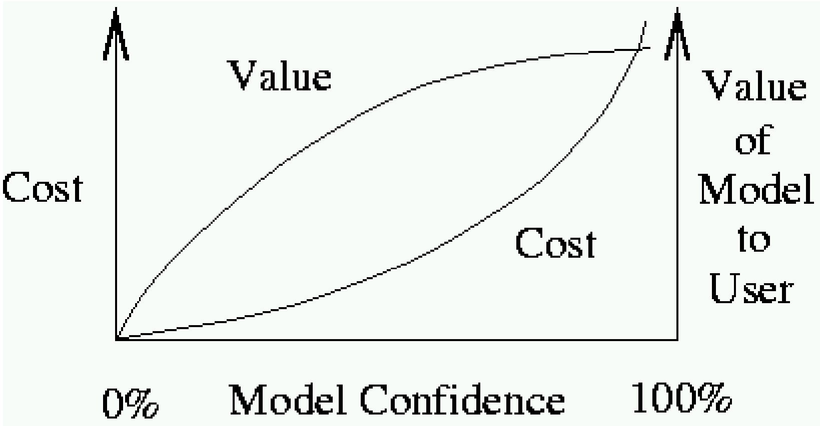

Figure 3 Cost and value to the user of a simulation system (Sargent, 2011)

1. An agreement be made prior to developing the model between (a) the model development team and (b) the model sponsors and (if possible) the users that specifies the basic validation approach and a minimum set of specific validation techniques to be used in the validation process.
2. Specify the amount of accuracy required of the simulation model's output variables of interest for the model's intended application prior to starting the development of the model or very early in the model development process.
3. Test, wherever possible, the assumptions and theories underlying the simulation model.
4. In each model iteration, perform at least face validity[1] on the conceptual model.
5. In each model iteration, at least explore the simulation model's behaviour using the computerized model.
6. In at least the last model iteration, make comparisons, if possible, between the simulation model and system behavior (output) data for at least a few sets of experimental conditions, and preferably for several sets.
7. Develop validation documentation for inclusion in the simulation model documentation.
8. If the simulation model is to be used over a period of time, develop a schedule for periodic review of the model's validity.

[1] Face validity: The extent to which a test is subjectively viewed as covering the concept it purports to measure (http://en.wikipedia.org/wiki/Face_validity (visited 2014.02.13))

3 Verification and validation work done by major users of simulation tools

In other industries, the focus on validation of simulation models have been given high priority as the consequences of basing designs or operational procedures on incomplete and/or faulty simulation models could be catastrophic. The project partners were asked to check out lessons learned and validation methods used by other transport areas such as air, rail and road. Some background papers have been uploaded to the project eRoom by Flanders Hydraulics Research and Ghent University. No feedback has been received from other partners.

MARINTEK has investigated validation processes used at the SINTEF's (and NTNU's) driving simulator (*http://www.sintef.no/home/Technology-and-Society/About-us/Transport-Research/laboratories/driving-simulator/).* Engen (Engen , 2008) has in his Doctoral thesis given a good summary of validation methods used for driving simulators, both behavioural and physical validation. He refers to another study (Kaptein et al., 1996) where it is stated that it is meaningless to define the validity of a research instrument. Validity should be defined with respect to a specific research question. Further Engen refers to Godley (Godley, 1999) and his two levels of validity:

- Behavioural validity or predictive validity, which looks at the correspondence between the real world and the simulator, in the way the human operator behaves
- Physical validity, looking at the physical correspondence between the simulator and its real world counterpart

The main topic of Engen's thesis is linked to behavioural validation, but he comments that it is important to undertake a physical validation study. This is especially true when the focus of the study is to examine vehicle handling. Two references to physical validation of a driving simulator are given in the thesis (Chrstos and Grygier, 1997) and (Salaani and Heydinger, 2000). Finally, a reference is given to a research project studying test participant's view of the ordinary car in the SINTEF/NTNU driving simulator (Moe, 1995). 17 professional drivers and 20 ordinary drivers took part in this study. The study was based on a questionnaire which was completed by the test persons after they had tried the simulator. The validation was divided into physical, operational and physiological experience related to the perceived realism of the simulator. The physical realism was considered high as a complete vehicle was used with all the equipment found in a real car. On a scale from 1 to 7 (1 – not realistic, 7 – very realistic) the average scores for operational realism were 3 from professional drivers and 4 from ordinary drivers, while physiological realism scored between 2.5 – 3 for professional and just about 4 for ordinary drivers.

Simulators are getting a strong position in the medical profession. A separate European society for simulation applied to medicine exists (http://www.sesam-web.org/). MARINTEK has discussed validation questions with Medisinsk SimulatorSenter (MSS, a joint centre between NTNU's Faculty of Medicine and St. Olav University Hospital) *http://www.simulatorsenteret.no/*. They apply an advanced instrumented and data controlled doll to create medical situations where a team of medical personnel has to react fast and apply the correct procedure to stabilize a life- threatening condition. Physical validation of the pre-programmed medical conditions has not been a topic at MSS as this should be handled by the doll manufacturer (Laerdal Medical). A paper describing use of simulators for nurse training (Elfrink Cordi et al., 2012) describes the lack of reliable and valid methods for measurement of human patient simulation effectiveness and gives an introduction to a Simulation Effectiveness Tool.

It should also be mentioned that there are a number of other simulators in use by the medical profession. In Trondheim there has been close cooperation between SINTEF/NTNU researchers and St. Olav medical staff in development of ultrasound equipment and training simulators. Virtual reality simulators are used for training of psychomotor laparoscopic skills. For such training it has been found that haptic sensations influence the psychomotor performance. A study has been published (Våpenstad et al., 2013) on how limitations of haptic feedback devices influence the construct validity of the applied virtual reality simulator. A more generic look at validation methodology for surgical simulation was published in 2005 (Aucar et al., 2005).

It is recommended that the validation processes used for airplane and train development and operations are reviewed as part of the project. If time permits it would also be beneficial to review process simulators (for instance those used in the oil and gas industry or nuclear power plants) and medical simulators should be checked and briefly described. This could be topics for student work.

4 Review of ITTC work on manoeuvring simulation and co-operative test programs

This section includes a review of findings from a study of ITTC Manoeuvring Committee reports. All reports can be found at http://ittc.info/download/proceedings.

Some ITTC members took part in the MARINER class co-operative test program, see Cummins (1969). Free-sailing and captive model tests were made with scaled models of a MARINER class vessel. Outcomes of these studies were used to predict full-scale manoeuvring performance and compare them with trials data for the case vessel "Compass Island". In the report from the Manoeuvring Committee it is stated:

- "Despite the urgent appeal for information in this field limited data is available, made up of a small number of relatively fine forms from Haslar, some Japanese data on cargo liners and tankers and some French data on car ferries and a four-screw liner".

In this report the final analysis of the 1st phase of the ITTC standard captive model test programme is presented and a short paper on correlation between model tests and ship trials data were given by A. J. Vosper.

The report from the Manoeuvring Committee to the 13th ITTC (Gertler and Norrbin, 1972) describes in brief some of the outcomes of the cooperative "Mariner" test program where different research organisations shared results from free-sailing and some captive model tests for the "Mariner" class vessel. From the tests numerical models were developed and simulation results compared with data from sea trials as one activity in investigating model –ship correlation. The report contains appendices on:

- Further analysis on model—ship correlation data
- Cooperative test program – review and status of 2nd phase of standard captive model test program

One of the recommendations from the committee was:

- Further experimental and theoretical information is required in support of a study of validity of the quasi-steady prediction methods now in use.

In the report from the 14th ITTC Manoeuvring Committee (Norrbin, 1975) it is stated that some additional results from the MARINER cooperative test program had been published since the previous meeting. It also includes a presentation of the ITTC 1975 Manoeuvring Trial Code to be used for comparative studies with free-running models. The following appendices are relevant for this MAROFF KPN project:

- Ship model correlation data (p. 373-403)
- Studies into the validity of quasi steady prediction techniques (p. 404-407)

At the 15th ITTC (van Manen, 1978) the Manoeuvring Committee in their review of relevant research activities mentioned the sea trials done in August 1977 with the VLCC "Esso Osaka". The study was sponsored by U. S. Coast Guard, Maritime Administration and American Institute of Merchant Shipping and included standard tests in deep, medium and shallow water. Trajectory data were measured and analysed by DTNSRDC, Washington.

Miller (1980) gave a status report on a correlation study using "Esso Osaka" as a case vessel. HYDRONAUTICS used a Large Amplitude Horizontal Planar Motion Mechanism to generate data for their simulation model. For the tests shown in the status report, there is rather good agreement between simulated and measured trials.

At the 16th ITTC (ITTC, 1981) the Manoeuvring Committee proposed that studies of "Esso Osaka" could be the subject of a new ITTC cooperative program where members did free-sailing and captive tests and use the sea trial data as a the base line for comparing simulated ship performance. One of the recommendations from the Manoeuvring Committee was:

- Ship model correlation techniques applicable to manoeuvrability studies should be kept under review. It is recommended that ITTC Member Organizations use their facilities to conduct model tests and compare their results with the "Esso Osaka" deep and shallow water data.

In the Manoeuvring Committee's report to 17th ITTC (ITTC, 1984) it is mentioned that a number of member organizations (19) have carried out model tests with "Esso Osaka" and compared outcomes to sea trial results. More work is needed to develop methods for model-ship correlation. In addition to the ITTC International Cooperation Program a Joint International Manoeuvring Program was initiated by the Maritime Administration, U. S. Department of Transportation and an ad hoc working group in Japan. Important papers on "Esso Osaka" studies have been written by Crane (Crane, 1979) and Goodman (1980a, 1980b).

In the report of the 18th ITTC the members are encouraged to continue to do model tests of the Esso Osaka tanker and to compare results with deep and shallow water sea trials. It is also mentioned that CETENA has a data bank for deep water sea trials including more than 100 vessels of different types and sizes (Chilo et al., 1984). Tests reported are turning circles, zig-zags and crash stops. Some more ships (30) have been reported by the Manoeuvring Committee (Wan, 1986).
One of the recommendations on future work from the 18th ITTC Manoeuvring Committee was:

- The subject of model and full scale correlation should continue to receive careful consideration.

For the 19th ITTC the Manoeuvring Committee recommended that member organisations continued to produce data from full scale ship manoeuvring trials for ship/model correlation. A number of papers for the Manoeuvring Committee were related to model test results and predictions of sea trials for "Esso Osaka". The report mentions that the U. S. National Research Council initiated to create a Committee for the Assessment of Shiphandling Simulation. In "Statement of task" for this committee it is stated (ITTC, 1990, p. 409) that:

- The accuracy of shiphandling simulators in all but the simplest shiphandling situations is not well known, nor are the accuracy requirements always appreciated.
- The confidence level of the results is difficult to establish.

A critical review of the validity of hydrodynamic models was thus the main task for this committee.

In the report of the 20th ITTC Manoeuvring Committee it is mentioned that the Committee developed and distributed a questionnaire to designers and builders. Two of the questions asked were:

- Have you any experience of scale effects in model manoeuvring predictions? If so, what form did they take and were they a problem?
- Do you use manoeuvring simulation as a training and/or design tool? If so, how was it validated?

30 replies were received and the findings were:

- Scale effects were experienced and corrections made using commonly used methods for friction resistance and wake fraction. Corrections were mainly focused on the longitudinal hydrodynamic force acting on the ship. Some also made corrections for propeller flow-straightening effect and shaft rpm. Results from scaling effects could be seen as a reduction in course stability with increasing ship length, in some cases the ship will be less course stable than the model.

- In most cases validation was done using full-scale results available from similar ships or using available model test data.

Section II.3 gives a summary on work related to scale effects, validation and accuracy. It is stated that the best way of validating physical or numerical model predictions is to compare predictions with outcomes from properly conducted sea trials. In this context it is mentioned that the major error source in sea trial measurements are the position and speed measurements. Use of GPS and DGPS is discussed. A reference is made to the ITTC 1975 Code of Practice for manoeuvring trials (ITTC, 1975) and it is proposed to add the following part into this code:

- An uncertainty analysis should be carried out on the completion of all trials, and estimates should be given of the accuracy of all measurements.

The committee proposes the accuracies for sea trial measurements given in Table 1.

Table 1 Proposed accuracy for sea trial measurements

Parameter	Value
Position of ship	± 3% of turning circle diameter D
Instantaneous speed of ship	± 3% of trial speed U
Propeller speed	± 1 rpm
Heading	± 0.5°
Heel angle of turn	± 0.5 °
Rudder angle	± 0.5 °
Rate of turn	± 3% of 360 x U/(π D) °/sec
Rudder stock torque	± 3% of maximum
Propeller shaft torque	± 3% of maximum
Propeller pitch	± 1% of maximum

According to the report most validation studies compare numerical and experimental results or physical model and sea trials, but do not analyse or quantify the errors. Knowledge of the errors is essential for all types of validation work. Dand (Dand 1992) discussed what could be meant by agreement – should it be complete correlation between prediction and measurements or could it be an overlap of the error bands (95% confidence limits)? In the context of simulator validation two methods are mainly used:

- Validation by pilots or experienced masters

- Validation against sea trials and/or model test results

Finally, the Committee discusses what should be meant by "good agreement" for simulation models. They recommend to relate this to the overlap between the error bands between prediction and those representative of "reality" (sea trials or free-sailing models). The need for coming up with practical validation criteria will be important and the 21st Manoeuvring Committee is requested to promote and evaluate relevant uncertainty analysis and benchmark data. N. H. Norrbin comments use of air fans during free-sailing manoeuvring tests. His statement is that he questions that using air fans gives the correct propeller loading and rudder force as long as there is no correction applied to the wake flow velocity. In a reply to a written question by J. Hooft the Manoeuvring Committee suggests that the new committee should consider an additional recommendation:

- Attempts should be made to initiate and simulate the development of a standard modular model for ship manoeuvring simulation together with unified methods for its validation.

In the report to the 21st ITTC Manoeuvring Committee reference is made to a MARIN Cooperative Research project looking at single screw displacement ships. The manoeuvring prediction program was validated using about 30 ships. The prediction of the 10/10 zig-zag was poor. In sec. 4.4 it is noted that it is necessary to rely on mariners in the final judge of validity or fidelity of simulation models. They refer to Webster and Young (1993) stating that:

- Validation is performed subjectively, often by pilots themselves, but preferably by a selected team of experts…. who observe and evaluate the simulation.
- Validation is usually an interactive process in which aspects of the mathematical model and physical components of the display are refined until consensus is reached.

In sec. 6 it is stated that validation of manoeuvrability predictions depends on the availability of systematic and reliable full scale trials. It is mentioned that very few full scale manoeuvring trials are carried out at a scientific level that make them suitable for validation studies. Further it is pointed to the fact that most trials data are for ballast or light load conditions while complex and challenging manoeuvres often are done at full load condition. Such manoeuvres are often done in confined waters while trials are done in deep, open waters. The Committee concludes that there is need for more accurate full scale data, preferably at full load condition, for validation purposes.

Tasks assigned to the 22nd ITTC Manoeuvring Committee included an activity to identify requirements for new procedures, benchmark data, validation and uncertainty analysis. A brief description of the extension of the MARIN Cooperative Research project to include twin-screw vessels is given and the outcomes of a new validation activity including more than 30 vessels point to differences between simulation model results and full scale data. Scale effects are according to the Committee mainly caused by the lack of similarity for velocity field in the stern region, especially over the rudder. In general, it is found that free-running model tests are more stable in connection with course keeping ability than the full scale ship. It is noted that a propeller loading between model and ship self-propulsion points gives the best agreement with full-scale results. Figure 4 illustrates a number of aspects in validation of simulation models. In discussion of benchmark data it is noted that very few full scale trials are made at a scientific level. One exception is the Esso Osaka trial. Chapter 8 of the report presents the outcome of the benchmark study initiated by the Manoeuvring Committee. 49 institutions responded to the survey, 17 sets of force data were available for the study. In the list of technical conclusions the Committee states:

- Validation of the hydrodynamic coefficients used in simulation models requires benchmark data at model scale.

- Validation of simulations, and of free-running model test results, requires benchmark data at full scale.
- Research is required to qualify scale effects on hydrodynamic forces acting on a manoeuvring vessel, particularly the lateral forces and yawing moments.
- Clarification is required on the scale effects present in free-running models.
- The available data for Esso Osaka in its current form is not suitable for use as benchmark.

Figure 4. Aspects for validation procedure for prediction of trials manoeuvres (ITTC 22nd Manoeuvring Committee, 1999)

Recommendation for future work includes a statement that the lack of benchmark data for all manoeuvring problems needs to be addressed by conducting suitable (free-running and captive) benchmark quality tests at model and full scale.

In section 6.2 of the report from the 23rd ITTC Manoeuvring Committee it is stated that effective validation for a ship handling simulator can be done by direct comparison of full scale and corresponding simulated ship manoeuvres. Further it is stated that validation of a complete modelling procedure is done by:

- Generation of hydrodynamic forces and moments from captive model test data, to construct a mathematical model
- Comparison of simulated trajectories from the mathematical model with outcomes from free-sailing model test data
- Further comparison of simulated trajectories from the mathematical model with full scale trial results

In the report it is noted that the Manoeuvring Trial Code has been rewritten and that a new procedure has been prepared for validation of manoeuvring simulation models. It outlines the general philosophy, but several issues remain unsolved in this draft procedure.

One of the tasks done by the 24th ITTC Manoeuvring Committee was to initiate the organization of the "Workshop on Verification and Validation of Ship Manoeuvring Simulation Models" (SIMMAN 2008, *http://www.simman2008.dk*). The Committee also worked on uncertainty analysis. It also states that it has been impossible to find relevant benchmark cases for manoeuvring. In section 4.2 the basic requirements for a benchmark case are given as:

- Full ship documentation (lines, propeller, engine, superstructure).
- Accurate full scale trial results, including full documentation of the ships loading condition.

The report reviews available benchmark ships and focus on the VLCC Esso Osaka, the Mariner ship and another VLCC (Esso Bernicia, reported by Clarke et al., 1972). The Committee has not been able to suggest a new benchmark ship that fulfils all requirements and chose to concentrate on manoeuvring prediction in model scale and propose to use the Japanese SR-221 tanker models as a model scale benchmark. Reference is also made to the models of a tanker made by the Korean towing tank KRISO. These models are part of the benchmark vessels defined for SIMMAN 2008. The other vessels are a container vessel (3600 TEU PanMax, L x B x T = 230 x 32.2 x 10.8 m, CB =0.651) and a navy vessel (DDG 51 frigate). The Committee ended up by recommending these 4 vessels as new benchmark ships for model scale manoeuvring studies. None of these designs will be built so there will be no full scale trial results for validation purposes. In the section on QM procedures it is stated that the Committee is responsible for 5 QM procedures:

- 7.5-02-05-05 Manoeuvrability Evaluation and Documentation HSMV.
- 7.5-02-06-01 Free Model Test.
- 7.5-02-06-02 Captive Model Test.
- 7.5-02-06-03 Validation Manoeuvring Simulation Models.
- 7.5-04-02-01 Full Scale Manoeuvring Trials.

In discussion of these procedures, it is commented that more work is needed with respect to uncertainty analysis.

The 25th ITTC Manoeuvring Committee presented an overview of accuracy and cost for different methods for documenting manoeuvring performance, see figure 5. Pros and cons for the different methods are discussed. More work is needed for validation of mathematical models, especially for non-standard manoeuvres at low speed or in shallow water. Section 5 of the report discusses the SIMMAN 2008 workshop and shows plots comparing predictions with mathematical models and free running model tests. Two of the general conclusions from the workshop are:

- There is a general need for more quantitative verification and validation.
- There is a general need for a definition of how to validate a manoeuvring prediction method, i.e. which accuracy is acceptable? If possible, a "prediction quality index" should be defined.

In the section discussing standards and safety they refer to a SNAME H-10 panel survey resulting in a recommendation of slow speed tests for documenting ship's performance. Further there are miner updates to most of the QM procedures. However, the procedure on Full Scale Manoeuvring Trials has been rewritten. Work has been done for a new procedure on Uncertainty Analysis of captive model tests. In their conclusions the Committee points to:

- Not much research is reported on scale effects for predictions based on model scale results.
- Validation and documentation is needed for mathematical models used in ship-handling. simulators, especially regarding non-standard manoeuvres, e.g. at slow speed and in shallow waters.
- Further research is required to develop a standard method for the correction os sea trial data obtained at non-standard trial conditions.
- The necessity of low-speed manoeuvring criteria has been raised. Some low-speed manoeuvres and indices have been proposed. For practical and meaningful low-speed criteria, it is required to obtain more information on that from both pilots and ship designers and to collect full scale manoeuvring data in the harbour and waterways.

Finally, ITTC is requested to adopt the procedure on UA in captive model tests named "Forces and Moments UA example for PMM tests".

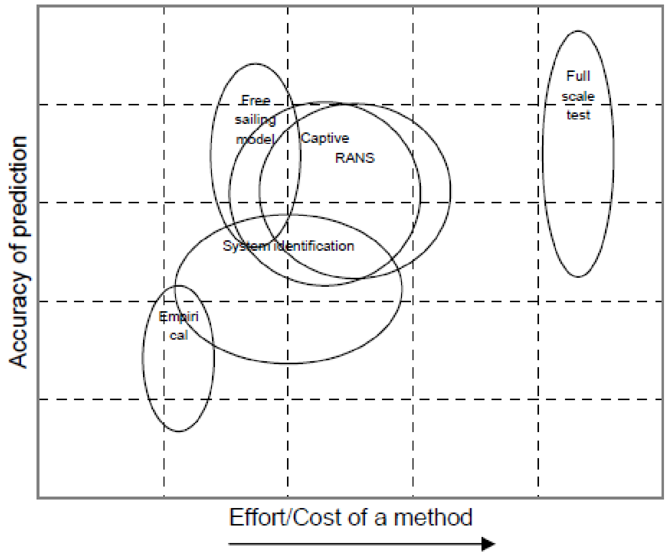

Figure 5 Effort/cost versus accuracy of manoeuvring prediction methods (Manoeuvring Committee, 25th ITTC, 2008)

The 26th ITTC took place in Rio de Janeiro in 2011. The Manoeuvring Committee report includes a section on benchmark data where the focus is on the work done in connection with the SIMMAN 2008 workshop and the 4 benchmark vessels KVLCC1 and KVLCC2 (tankers), KCS container vessel and 5415 navy vessel. In a comment to the findings from captive model tests it is mentioned that different institutions perform the tests with different propulsion parameters. INSEAN ran the tests at model self-propulsion point using a constant torque while the MOERI tests were done at ship self-propulsion point and constant RPM. It is recommended to continue work to establish a full set of well-documented benchmark model test data as well as corresponding free sailing tests with the SIMMAN vessels.

In a discussion of uncertainty analysis it is stated that a worked-out example for captive model testing is a prerequisite for transfer of the existing procedure towards the recommended ISO approach. It was recommended that "Knowledge on the uncertainty of the outcome of the mathematical model as a whole" should be the "leading star" for the 27th Manoeuvring Committee.

In the Committee's discussion of scale effects, they point to the fact that there is no established standard test and correction method to overcome scale effects. Reliable model test and full-scale trial data are required. The report briefly describes scale effects on rudder forces, hull forces and propeller and rudder wake field and flow straightening. It is mentioned that institutions have varying practice for performing manoeuvring tests. Some use model self propulsion point others use ship self propulsion point. In their concluding remark on scale effects the committee states that some recent studies indicate that the present prediction model based on free or captive model tests may fail to predict full-scale manoeuvring performance. This is especially critical for unconventional vessels with new propulsion systems. A guideline for how to select optimum propulsion point in model scale to get the best full scale prediction is needed.

The Committee presents a brief discussion on manoeuvring models for low-speed operation. The need for model validation on both force level and motion level require that a series of low-speed manoeuvres are selected and standardised and used for validation studies.

The 27th ITTC Manoeuvring Committee met in Copenhagen 31August – 6 September 2014 (ITTC, 2014). The document describing tasks for all ITTC committees has a number of general terms of reference. Here it is stated that all new procedures for uncertainty analysis in experiments shall follow the ISO "Guide to the expression of Uncertainty in Measurements" (ISO, 1995). In the special terms of reference for the Manoeuvring Committee the following statements are relevant for this MAROFF project:

- 1f – The need for R&D for improving methods for model experiments, numerical modelling and full-scale measurements.
- 3. – Complete the work on the Procedure 7.5-02-06-04 Uncertainty Analysis; Forces and Moment, Example for Planar Motion Mechanism Test, based on ISO approach.
- 6a – Update the procedure 7.5-02-06-01 Free Running Model Test Procedure
- 7 – Report on knowledge and collect, analyse and summarize data on scale effects for manoeuvring predictions.
- 10a – Report on manoeuvring criteria for ships not directly covered by IMO like POD and waterjet driven vessels, naval vessels, inland ships, HSMV, etc.
- 10b – study possible criteria for manoeuvring at low speed and if warranted communicate findings to IMO.

The main tasks completed by the 27th Manoeuvring Committee were a (small) update of the procedure of validation of manoeuvring simulation methods, the creation of the guidelines on uncertainty analysis for free running model tests, on V&V of CFD in manoeuvring, a discussion on manoeuvring criteria and scale effects. According to Guillaume Delafortrie (member of the ITTC Manoeuvring Committee) the terms of reference for the 28th ITTC Manoeuvring Committee which are of interest for the SIMVAL project are:

- 1f. the effects of free surface, roll, sinkage, heel and trim in numerical simulation of manoeuvring one example is for instance found in the observed differences for the KCS at the SIMMAN2014 workshop, we are still waiting for the conclusions of that workshop, but it will be certainly of interest for the project.
- 4. Update ITTC Recommended Procedure 7.5-04-02-01, Full Scale Manoeuvring Trials Procedure. Include consideration of full scale to model scale correlation. In particular, examine the model scale to full scale correlation of steering control in manoeuvring.
    - The procedure is from 2002, so it may certainly need an update, as the IMO Manoeuvring Criteria came into charge in 2002. Scale effects, rudder control and use of autopilots are also issues determined by our Committee. We also need to check the corresponding ISO procedure.
- 6. Continue work in order to have a full set of benchmark data for each of the benchmark hulls (KVLCC2, KCS, 5415, HTC, SUBOFF and S175 – manoeuvring in waves).
- 7. Extend the uncertainty analysis for captive model tests from measurements towards a new procedure that provides the uncertainty analysis for the use of captive model tests in the predictions of manoeuvring. Elaborate with an example.
    - It will probably be a kind of sensitivity analysis as used for the uncertainty analysis of free running tests. The main issue is the uncertainty assessment of filtering of measurements.
- 9.Conduct a concise review and report on the specific aspects of the manoeuvring of vessels in restricted waters such as ports and harbours.
- 10.Review testing methods for ships in ports and harbours, including ship mooring loads, safe speed limits for moving and its impact on moored vessels, manoeuvring concerns such as squat and bank effects.

5 Review of previous manoeuvring validation studies

SNAME H-10 panel initiated studies on validation of manoeuvring simulation for selected vessel types through studies of standard manoeuvres in the 1960/70-ties. Studies were made using models of Mariner class vessels (a general type cargo vessel). Outcomes of these studies are described in reports from ITTC's Manoeuvring Committee in 1972 and 1975.

Exxon funded a major manoeuvring study for their VLCC "Esso Osaka" in the late 1979-ties/early 1980-ties. The vessel performed standard manoeuvres in three different water depths given by 20%, 50% and 320% under-keel clearances (UKC). Data from the tests have been made public in reports by Lincoln Crane Jr. et al (1979).

In the early 1990-ties the US National Research Council appointed the "Committee on Assessment of Shiphandling Simulation". They prepared a report (Webster et al, 1992) on application of shiphandling simulation to waterway design. Here they state in Ch. 6 that it is impossible to assess scientifically the accuracy of either the mathematical model or the pilotage performance of individuals. Instead, validity is obtained if pilots conclude that it accurately reproduces the modelled ship's behaviour in a particular waterway. They apply the idea that if several pilots familiar with the waterway, the modelled ship or both are all satisfied with the simulation, there is a reasonable confidence in the results. They point to validation as being an iterative process. Application of this approach is not feasible for new ship designs. In Appendix F of the report a short presentation of validation for aircraft flight simulators is given. Here it is noted that simulators for aircraft are validated following guidelines prepared by aviation authorities and that simulation software for military systems is designed, documented and validated against a specific standard (DOD-STD-2167). For validation of an aircraft flight simulator one usually goes through four levels of testing:

- Individual tests of modules or sub-programs.

- Testing of small program packages or groups of sub-programs that are related functionally.
- In level 3 the complete mathematical model is tested for dynamic response verification.
- The final level consists of pilot-in-the-loop testing.

The SIMMAN 2008 workshop gathered manoeuvring experts from research companies and universities in an attempt to study reliability of existing methods for studying ship manoeuvring performance. Lines were made available for selected vessels:

- Two different versions of a VLCC
- A container vessel
- A naval vessel

Research organisations performed captive and free-sailing models tests, developed simulation models based on model test results or empirical expressions. Results from model tests and numerical simulations were used in a "blind test" where the base case was a set of standard manoeuvres found from free-sailing test. IMO's standard manoeuvres were used as case manoeuvres. MARINTEK's contribution to SIMMAN was based on calculations for the two VLCC versions.

In December 2014 the second SIMMAN workshop (SIMMAN 2014) took place in Copenhagen. At this workshop studies of low speed performance and shallow water influence were studied. NTNU's PhD candidate Sergey Gavrilin took part in the workshop – in his summary from the participation he highlighted:

- The main goal of the workshop was to compare predictions of standard IMO trials done by different models. Compared to previous SIMMAN workshop, which was held in 2008, only three case vessels were considered (KVLCC2, KCS and 5415).
- However, the test program was extended by inclusion of shallow water trials. New captive and free running model tests in deep and shallow water were conducted prior to the workshop.
- Submissions made by participants were compared and presented in ten sessions: *EFD-FRMT, Trajectories, Math models, System based forces, EFD – Captive, CFD forces, CFD flowfield, Forces in time, Shallow Water and Conclusion*.
- The main attention was paid to models based on CFD. CFD methods were mainly applied for virtual captive tests. However, some free running model tests simulated by CFD were presented as well.

MARINTEK and NTNU did not present any results at SIMMAN 2014. Flanders Hydraulics Research had performed the shallow water PMM tests for the KVLCC2 tanker and Katrien Eloot was chairman of the Shallow water session. NRNU's PhD candidate Sergey Gavrilin t

6 Existing guidelines for validation of ship manoeuvring simulation models

ITTC updated in 2011 their procedure 7.5-02 06-03, rev. 2, (ITTC 2011b). In the validation section it is stated that the applied prediction method must be validated against benchmark data. The documentation of such validation should be available in form of a report or published paper and should include the outcomes of 6 activities:

- Prediction of hydrodynamic forces
- Modelling of forces in a mathematical framework
- Mathematical model structure
- Integration model
- Simulation software
- Simulated manoeuvres

6.1 Evaluation of existing guidelines

As shown in the review of previous work on validation of simulation models for ships, there is a lack of high quality sea trial data for model validation work. The ongoing R&D work for the coming SIMMAN 2014 will focus on selected basic design vessels for which there will be no sea trial data. For validation purposes the outcome of free-sailing models will be used. The MAROFF project has been designed to establish a database of sea trials for selected case vessels. These data should be used for validation purposes.
The ITTC procedure 7.5-02 06-03, rev.2 should be used as a baseline for documentation of validation studies for numerical simulation models. Definitions of the terms verification, validation and accreditation (VV&A) should as far as possible follow the new definitions that will be included in the 2014 version of the ITTC Terms of Reference.

6.2 Validation session at OMAE 2015

A separate session on ship simulation model verification will be arranged during the Ocean Engineering symposium at OMAE on June 1st 2015. Under the topic

6-1 Advanced Ship Hydromechanics and Marine Technology

the sub-sessions

6-1-1 Marine simulation models and their validation techniques

6-1-4 Marine Simulation Models II/Canadian Shipbuilding III

includes the following papers related to SIMVAL activities.

- Uncertainty of Sea Trials Results Used for Validation of Ship Manoeuvring Simulation Models, authors Sergey Gavrilin and Sverre Steen
- Validation of ship manoeuvring in shallow water through free-running tests, authors Katrien Eloot, Guillaume Delefortrie, Marc Vantorre and Frans Quadvlieg
- Validation of a Modular Mathematical Model for Low-Speed Maneuvering using Small Scale Tests with an Oceanographic Research Vessel, authors Felipe R. Masetti, Pedro C. de Mello and Eduardo A. Tannuri
- Time Domain Simulation Model for Research Vessel Gunnerus, authors Vahid Hassani, Ørjan Selvik, Andrew Ross, Dariusz Fathi, Florian Sprenger and Tor Einar Berg
- Identification of Nonlinear Manoeuvring Models for Marine Vessels using Planar Motion Mechanism Tests, authors Andrew Ross, Vahid Hassani, Ørjan Selvik, Edvard Ringen and Dariusz Fathi

7 Recommended validation guideline to be used in this project

It is proposed that the scientific partners of the project initially try to apply the 8 steps given by Sargent (see p. 4-5) for the KVLCC2 model (being part of the SIMMAN 2014 study) as benchmark data to get familiar with this verification and validation method. Doing this one should look at the statements on the validation process found in the ITTC procedure 7.5-02 06-03, rev.2. Only the deep water part of ship manoeuvring should be investigated in this initial activity. The research partners should discuss the need for and possible outcomes from an uncertainty analysis following the guidelines prepared by ITTC.

It is recommended that the scientific partners of the project meet in connection with the May 2015 Steering Committee meeting to review new information on mathematical model verification and validation tabled by the 27th Manoeuvring Committee of ITTC.

It is further recommended that the scientific partners studies and comments the verification and validation work which will be done by MARINTEK and NTNU using the NTNU research vessel "Gunnerus". For this vessel, the line plan will be available for all research partners. In addition, captive model test results, free-sailing tests and sea trial data will be available for this vessel. A workshop will be arranged in connection with the Steering Committee meeting in May 2015.

8 Conclusions

The study has shown that validation of mathematical models for shiphandling has a long way to go. Very few cases of validation studies using sea trial data have been published. As has been found for Esso Osaka investigations, the quality of sea trial data has a lot of uncertainties.

MARINTEK's experience with manoeuvring sea trial data collected by the shipyard in connection with delivery trials, are that they are unreliable. In general, they do not have the quality needed for validation studies of advanced mathematical models for ship manoeuvring and seakeeping.

For the validation studies it will be important to discuss validation in the context of how the simulation models will be applied. The validity level needed for training of generic shiphandling skills for different standard ship types and sizes are far lower than that needed for vessel and operation specific shiphandling skills for unconventional vessels.

The scientific partners should prepare text proposals for master thesis topics on simulation model verification and validation and distribute them to new master students.

A References

Aucar, J. A, Gronch, N. R., Troxel, S. A. and Eubanks, S. W.: A review of surgical simulation with attention to validation methodology. Surg Laparosc Endosc Percutan Tech – volume 15, number 2, pp. 82-89, April 2005

Bair, L. J.: The missing link in modeling and simulation validation, 2012 Autumn Simulation Multi-Conference, San Diego, October 2012

Berg, T.E and Ringen, E.: Validation of shiphandling simulation models. Proceedings of the 30th International Conference on Ocean, Offshore and Arctic Engineering, paper no. OMAE 2011-50107, Rotterdam, The Netherlands, June 2011

Chilo, B., Della Loggia, B., Deutsch, C. and Routin, M.: Ship manoeuvrability aspects and human factors. Proc. International Conference on Aids to Safe Maritime Traffic in Mediterranean Areas, Naples, 1984.

Chrstos, J. P., and Grygier, P. A. (1997). "Experimental testing of a 1994 Ford Taurus for nadsdyna validation." SAE transactions, 106(6), 895-908.

Clarke, D., Patterson, D. R. and Wooderson, R. K.: Mnoeuvring trials with the 193.000 tonne deadweight tanker Esso Bernicia. RINA spring meeting. Paper no. 10, 1972.

Crane, C. L.: Maneuvering trials of the 278 000 dwt "ESSO OSAKA" in shallow and deep water. Transactions SNAME, Vol. 87, 1979.

Cummins, W. E.: Report of the Manoeuvring Committee. 12th International Towing Tank Conference. Rome, September 1969

Dand, I.W.: Measurement errors and manoeuvring predictions. Workshop on Prediction of Ship Manoeuvrability. The West-Japan Society of Naval Architects. Fukuoka, 1992.

Department of Defense, 1988: DOD-STD-2167: Military Standard: Defense system software development. United States Department of Defense, 29. February, 1988

Elfrink Cordi, V. L., Leighton, K., Ryan-Wenger, N., Doyle, T. J. and Ravert, P.: History and development of the Simulation Effectiveness Tool (SET). Clinical Simulation in Nursing, 8(6), e199-e210. Doi:10.1016/j.ecns.2011.12.001

Engen, T.: Use and validation of driving simulators. NTNU Doctoral thesis 2008:200, ISBN 978-82-471-1083-6. Trondheim, 2008

Gertler, M. and Norrbin, N. H.: Proceedings 13th International Towing Tank Conference. Berlin, 1972

Godley, S. T. (1999). "A driving simulator investigation of perceptual countermeasures to speeding." Monash University. Dept. of Psychology.

Goodman, A., 1980a: Comparison of results between computer simulations and full-scale trials for ESSO OSAKA. 16th ITCC Manoeuvring Committee working paper, 1980

Goodman, A., 1980b.: Model test and simulation correlation study based on the full scale "ESSO OSAKA" maneuvering data. U. S. Department of Commerce, Maritime Administration, 1980.

Harmon, S.Y. and Youngblood, S.A.: A proposed model for simulation validation process maturity, Journal of Defense Modeling and Simulation 2 (October 2005), no.4:179–190

IMO: Standards for ship manoeuvrability, Resolution MSC137(76), IMO 2002

ISO: Guide to the expression of Uncertainty in Measurements" (ISO, 1995)

ITTC 1975: Code of practice for manoeuvring trials, 14th International Towing Tank Conference. Ottawa, September 1975

ITTC 1981: 16th International Towing Tank Conference. Leningrad, August-September 1981(p. 298).

ITTC 1984: 17th International Towing Tank Conference. Gothenburg, September 1984

ITTC 1987: 18th International Towing Tank Conference. Kobe, October 1987

ITTC 1990: 19th International Towing Tank Conference. Madrid, October 1990

ITTC 1993: 20th International Towing Tank Conference. San Francisco, October 1993

ITTC 1996: 21st International Towing Tank Conference. Trondheim, October 1996

ITTC 1999: 22nd International Towing Tank Conference. Seoul and Shanghai, October 1999

ITTC 2002: 23rd International Towing Tank Conference. Venice, October 2002

ITTC 2005: 24th International Towing Tank Conference. Edinburgh, October 2005
ITTC 2008: 25th International Towing Tank Conference. Fukuoka, October 2008
ITTC 2011a: 26th International Towing Tank Conference. Rio de Janeiro, October 2011
ITTC 2011b: Validation of manoeuvring simulation models. ITTC – Recommended procedures and guidelines, 7.5-02 06-03, Revision 02, ITTC 2011
ITTC 2014: 27th International Towing Tank Conference. Copenhagen, September 2014
Kaptein, N. A., Theeuwes, J., and Van Der Horst, R. (1996). "Driving simulator validity: Some considerations." Transportation research record(1550), 30-36.
Miller, E. R, jr.: Status report on a model-full scale correlation study based on the "Esso Osaka" maneuvering data. In Proceedings of the Nineteenth General Meeting of the American Towing Tank Conference, Vol. II, p1237-1267, Ann Arbor, July 1980.
Moe, D. (1995). "Evaluering av bilsimulator Autosim A/S – del1: Alphatest." Sintef samferdselsteknikk, Trondheim (in Norwegian).
Norrbin, N. H.: Report of the Manoeuvring Committee. 14th International Towing Tank Conference. Ottawa, September 1975
Pace, D. K: Modeling and simulation verification and validation challenges. John Hopkins APL Technical Digest, Vol. 25, no. 2 (2004), p. 163-172, Maryland
Salaani, M. K., and Heydinger, G. J. (2000). "Model Validation of The 1997 Jeep Cherokee For the National Advanced Driving Simulator." 2000-01-0700.
Sargent, R.G.: Some Approaches and Paradigms for Verifying and Validating Simulation Models. In Proceedings 2001 Winter Simulation Conf., edited by B. A. Peters, J. S. Smith, D. J Medeiros, and M. W. Rohrer, 106-114. Piscataway, New Jersey: Institute of Electrical and Electronic Engineers Inc.
Sargent, R. G.: Verification and validation of simulation models. In Proceedings 2011 Winter Simulation Conference, edited by S. Jain, R.R. Creasey, J. Himmelspach, K.P. White and M. Fu.
Stern, F. and Agdrup, K.: SIMMAN 2008 – Workshop on Verification and Validation of Ship Manoeuvring Simulation Methods. Copenhagen 14th – 16th April 2008. FORCE Technology and Iowa Institute of Hydraulic Research, April 2008.
Van Manen, : Report of the Manoeuvring Committee. 15th International Towing Tank Conference. Wageningen, September 1978
Våpenstad, C., Fagertun Hofstad, E., Bø, L. E.,Chmarra, M. K., Kuhry, E., Johnsen, G., Mårvik, R. and Langø, T.: Limitations of haptic feedback devices on construct validity of the LapSim® virtual reality simulator. Surg Endosc (2013) 27:1386-1396, DOI 10.1007s00464-012-2621-9
Wan, T.: Summary of trial data on full-scale manoeuvring tests. China Ship Scientific Research Centre's Contribution, 1986.
Webster, W. (editor): Shiphandling simulation: Application to waterway design. National Research Council. ISBN: 0-309-59802-8, 1992
Webster, W. and Young, W.: Marine Board ship simulator research. Proceedings of MARSIM'93. International Conference on Marine Simulation and Ship Manoeuvrability, St. Johns, Canada, 1993

B Abbreviations

Abbreviation	Definition
CFD	Computational fluid dynamics
DGPS	Differential global positioning system
DOF	Degree of freedom
DTNSRDC	David W. Taylor Naval Ship Research & Development Center
GPS	Global positioning system
HSMV	High-speed marine vehicle
IMO	International Maritime Organisation
INSEAN	The Italian Ship Model Basin
ISO	International Standard Organisation
ITTC	International Towing Tank Conference
KPN	Knowledge-building Project for Industry
KRISO	Korea Research Institute of Ships and Ocean Engineering
MOERI	Maritime & Ocean Engineering Research Institute
NTNU	Norwegian University of Science and Technology
OMAE	Ocean Mechanics and Arctic Engineering
PMM	Planar motion mechanism
QM	Quality management
RPM	Revolutions per minute
R&D	Research and development
SIMMAN	The Workshop on Verification and Validation of Ship Manoeuvring Simulation Methods
SNAME	Society of Naval Architects and Marine Engineers
TEU	Twenty-foot equivalent unit
UA	Uncertainty analysis
UKC	Under keel clearance
US	United States
VLCC	Very large crude carrier
VPMM	Validation process maturity model
V&V	Verification and validation
VV&A	Verification, Validation and Accreditation