

Comparing Simulation with Physical Verification and Validation in a Maritime Test Field

Arnold Akkermann¹, Axel Hahn²

¹R&D-Division Transportation, Institute for Computer Sciences, Oldenburg, Germany

²Department for Computer Science, University of Oldenburg, Oldenburg, Germany

Email address:

Arnold.Akkermann@offis.de (A. Akkermann), axel.hahn@uol.de (A. Hahn)

To cite this article:

Arnold Akkermann, Axel Hahn. Comparing Simulation with Physical Verification and Validation in a Maritime Test Field- Opportunities and Challenges –. *International Journal of Systems Engineering*. Vol. 4, No. 2, 2020, pp. 18-29. doi: 10.11648/j.ijse.20200402.12

Received: September 8, 2020; **Accepted:** October 6, 2020; **Published:** October 21, 2020

Abstract: The steadily increasing complexity of maritime systems substantially raised the need for advanced verification and validation (V&V) as well as certification methods. Extensive simulation-based certification adds new opportunities to existing physical testing. Compared with simulation, field tests are extremely time-consuming and therefore expensive. Furthermore, relevant close-range situations between ships or environmental impacts (e.g. certain types of bad weather situation) are impossible to perform in the field for safety reasons and the uncontrollability of the environment or simply the amount of experiments needed. Systems in the maritime domain (like products for navigation assistance, sensors, communication equipment etc.) are typically not used isolated but as part of a complex setup. More and more sensors and actuators are integrated to provide data for various systems or information services on board a ship and ashore. Since such systems are typically continuously evolving during their service lifetime, the development and maintenance of maritime systems (e.g. bridge systems) need to be considered in its usage context that includes interconnected systems and external services, sensors and actuators. CPSoS (Cyber-Physical System of Systems) demand innovative approaches for distributed optimization, novel distributed management and control methodologies that can also deal with partially autonomous systems, and must be resilient to faults or cyber-attacks. In addition, CPSoS engineering no longer maintains the former strict separation between the engineering phases and actual operation. Instead, integrated approaches for the design- and operation- phase are required to cover the full lifecycle by modelling, simulation, validation, and verification (V&V). Thus, prospectively, it will be necessary to monitor the system formation and to conduct a final assessment of the system by means of a suitable application of test cases in a controlled and comprehensible manner. These systems have an emerging behavior and cannot entirely be defined during the design phase. At this point it becomes apparent that conventional unit, integration and system tests are no longer sufficient to fully cover and validate the functional limits of Cyber-Physical System of Systems. An acceptable test coverage cannot be achieved with these methods for such systems. In this paper the authors present a use case of collision-regulation compliance checker to compare virtual (i.e. simulation-based) V&V, physical (i.e. in-situ testing) V&V and hybrid, mixed-reality V&V.

Keywords: Scenario-based Testing, V&V Lab, Mobile and In-Situ Platform

1. Introduction

Simulation based testing and verification and validation is a crucial technology in the early phases of systems design. Further down the design process with the availability of prototypes physical tests are possible and provide the required grounding. Physical tests are much more expensive and critical situations potentially generate risks which are not bearable. Additionally, today's complex systems show nondeterministic behavior and need very extensive testing to

identify rare critical events. It would take thousands if not millions of test-runs to identify such rare events. What is the power and the expressiveness of virtual testing and is it possible to drastically reduce the number of physical tests and can virtual testing play a significant role in certification? What are the opportunities and challenges of virtual V&V (Validation and Verification)? As a use case this paper compares virtual and physical testing of a collision-regulation compliance checker as the system under test. The use case is performed with the generic maritime testbed

eMIR (eMaritime Integrated Reference Platform) which provides a seamless (V&V) approach from virtual, simulation-based testing to performing physical tests within a maritime open sea testbed that offers an extensive ship and shore-sided geographically distributed infrastructure. Simulation and in-situ testing already is a widely adapted practice for maritime system qualification, but eMIR seamlessly interconnects both worlds. Each test can be configured to be performed either based on a full-featured maritime traffic simulation platform, be interconnected with hardware components (i.e. hardware in the loop testing), or even be performed in-situ in our open sea testbed with vessels equipped with a second, container-based bridge. The simulation and the physical experiment both have the same interfaces to the system under test and therefore it is possible to switch back and forth between the simulated world and the concrete physical test field. Figure 1 gives an overview about the basic components of eMIR. The in-Situ platform is a region in the German Bight, which has instrumented by shore-based sensor systems Radar (Radio Detection and Ranging), video, weather, and AIS (Automatic Identification System) sensors). The platform is design for model building (e.g. traffic behavior patterns, close-by situations) and extends common vessel dynamic models to use for maritime traffic and maneuver simulation in the V&V Lab, which provides of a server-based infrastructure for Big Data analysis and several interconnected ship bridge and maritime traffic simulators. For in-Situ tests the relevant vessels are equipped with a complete second bridge system (ECIDS (Electronic Chart Display and Information System), Conning (information display system for symbol illustrating the configuration of the ship), and ARPA (Automatic Radar Plotting Aid), inside sea containers that come also with all relevant sensors such as GPS (Global Positioning System),

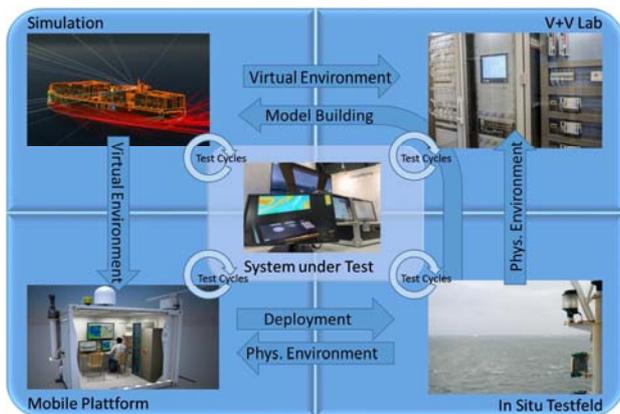


Figure 1. Components of the eMIR testbed.

Radar, AIS, LIDAR (Light Detection and Ranging), and CV (Camera Video) Systems or small bridge systems are embedded into transportable cases for onboard installation. The mobile platforms additionally can inter-connected with the shore-based infrastructure via satellite connections and standard NMEA- (National Marine Electronics Association) based systems of the vessels.

This paper has the following structure: The next section discusses relevant related work focusing on maritime test beds. Thereafter, in section 3, we describe the structure and basic concepts integrated in eMIR in detail. We present and discuss the traditional maritime system testing approach (i.e. a HAZOP (Hazard and Operability) analysis) followed by an overview about the four basic components of the testbed: Simulation, V&V Lab, the Mobile and the In-Situ Platform. Then we present a use cases in that we applied the ACTRESS (Architecture and Testbed for Realtime Safe and Secure Systems) approach for verification: the qualification and the verification of a component of a maritime Collision Avoidance System. The use case is describe using the following structure: (1) A brief description of the system under test, (2) the simulation part, (3) the in-situ testing, and (4) the outcome and lessons learned. Finally, we summarize our contribution and state future work.

2. Related Work

In the maritime domain, a number of test beds for new e-navigation and monitoring technologies are currently in planning and implementation worldwide. These test environments have different objectives: To understand the challenges and requirements for e-navigation, to develop and test (validate and verify) platforms or to demonstrate the maturity of new technologies. Testbeds have been developed in many maritime projects to test specific technologies, e.g. in the North Sea [1], the Baltic Sea [2-4], the Ionian Sea [5], the Adriatic Sea [6], the Malacca Strait [7] and Japan [8]. Each of these testbeds is specialized by its individual applications. In general, however, most testbeds of them focus on optimizing the planning and coordination of ship movements in order to increase safety at sea. A central role plays the exchange of information between ships and shore stations in order to derive new functions and improved functionality for each testbed at sea. Generic testbed platforms are another approach to testing new technologies: reusable and configurable. This approach adopts concepts from the automotive industry such as the application platform for intelligent mobility (AIM). AIM is a component-based testbed for land transport [9]. It has mobile components like a vehicle fleet or structural components like a research crossing, a research level crossing or the reference route with Car2X infrastructure. It also offers driving or virtual reality simulators. These components are test carriers for new technologies. The collected data can afterwards used for simulations or other research work. The development of such systems in the automotive sector is at a more advanced stage than in the maritime sector. Due to the highly dynamic market situation, a lot of effort being put into developing not only the automated driving functions but also the necessary V+V methods. However, the maritime domain is no less complex than the automotive sector. This domain characterized by a multitude of different ship classes and types (cargo ships, tankers, etc.) with specific propulsion systems (diesel engines, gas turbines, fixed pitch propellers, controllable pitch propellers). Each ship type has different maneuvering characteristics (stop

lengths, maneuvering behavior), and is operating on long shipping routes in different sea areas (shallow water, restricted waters, deep unrestricted waters) in different traffic situations (head on, overtaking, crossing) under changing weather conditions. While the system has much lower dynamics due to high inertia that comes with the need for extensive need for prediction during maneuvers. This diversity results in several thousand test cases, for example for testing the behavior of an autopilot, in order to achieve an acceptable test coverage without considering the specific requirements of the autopilot. Scenario-based testing offers an effective approach to reducing the number of test cases. Here, for example, the behavior of the autopilot in a specific traffic situation within a scenario can be tested and compared under different weather conditions (current, swell) for different types of ships with different propulsion systems. In addition, within the same scenario the behavior at different rudder positions at different speeds can be tested. This means that the number of test cases will effectively minimize through the parameterization of scenarios without reducing the test coverage. For that background, we have chosen this approach in our research work.

3. Scenario Based V+V with eMaritime Integrated Reference Platform

Connected and loosely coupled systems from different manufacturers characterized the development of automated ship systems. This leads to a steadily growing number of complex system developments and the usage of intensive verification and validation methods. Test fields are particularly suitable for supporting and performing verification and validation (V+V) during the entire system development process and offer the possibility to test the system in a simulative or physical environment. The heterogeneity of the systems to be tested and extensive requirements present challenges for the design of current maritime test setups. However, for the execution of given test scenarios, these setups are designed individually and are not viable because the design follows unstructured methods. In contrast, the eMIR platform implements a system architecture for a sustainable and reusable virtual and physical test field. EMIR provides a framework for engineering, validation, verification, and demonstration of technological innovations as for new cooperation and process models. This platform integrates the virtual testbed HAGGIS (Hybrid Architecture for Granularly, Generic and Interoperable Simulations) and the physical test field LABSKAUS (German: LABor für SicherheitsKritische Analysen aUf See / English: laboratory for safety critical analysis on sea).

3.1. Architecture and Components of the eMIR Testbed

To support the whole development process of highly automated and autonomous systems and in order to be able to meet future CPS requirements as well, a maritime testbed should be open and sustainable. Open testbed

means that eMIR is open to integrate new technologies, services or sub platforms. For comprehensive interoperability for SoS (System of Systems) under test and to implement shared services to facilitate CPSoS (Cyber-Physical System of Systems) testing, the following described testbed provides a shared infrastructure and interoperability architecture. In order to reflect the concept of interoperability, eMIR based on the uniform data exchange format S-100 as connection between all existing elements and components. In order to meet the requirements for a novel testbed for highly automated systems, a possibility must create to check the predefined use cases of the System under Test (SuT) using concrete test scenarios and thereby validate the correct functionality of the system under test. In the following, the interconnected physical and virtual testbed components will discuss in more detail.

3.1.1. Virtual / Simulation Based Verification and Validation Environment HAGGIS

HAGGIS is a modelling and open co-simulation environment to build virtual e-Navigation testbeds and is part of the virtual eMIR testbed. It enables rapid testing of new e-Navigation technologies in a simulation environment. HAGGIS consists of a number of modules that are orchestrate for different applications. These allow simulating sensors, traffic or environment. HAGGIS consists of the Maritime Traffic Simulation for implementing, executing and observing the behavior of multiple vessels. Further the Environment Simulation Component for environmental scenario generation with several environment layers, the Sensor Data Simulation Component to provide the simulation information in a sensor specific format, several behavior simulations for artificial generated vessels and the World Editor for providing a system model to allow the setting up of an initial scene according to a predefined scenario. To perform safety analysis on the simulated scenarios, it is possible to initiate a risk monitor that can determine predefined risk situations during simulation. Distributed Controlling Toolkit (DistriCT) is this component. To ensure the technical interoperability of the HAGGIS components, the High-Level Architecture (HLA) as the communication middleware is used.

3.1.2. LABSKAUS

The physical parts of the eMIR testbed LABSKAUS are located in the German Bight. Even though some components are transportable and can be located elsewhere. One essential component of the eMIR testbed is the Reference Waterway with communication and surveillance technology along the Elbe River lane estuary between Cuxhaven and Brunsbüttel; Germany. The Reference Waterway is equipped with common and up to date maritime sensors, by expandable sensor nodes, including compact sensor data hubs, which provides navigational data on board a ship as well as data for maritime surveillance systems, to observe the maritime traffic. The eMIR testbed also includes two experimental ship-bridges for different

use cases or user requirements. An Experimental Bridge can be used i. a. as a maritime control station to imitate situational awareness and V&V management tasks during a test run. These Experimental Bridges are installed in a 10-foot respectively 20-foot CSC-certified Container as a harmonized and transportable solution. The Mobile Bridge is a modular solution for e.g. bridge component tests. To control a modelled "own ship" in the HAGGIS simulation, it is designed as a versatile mobile bridge to control a research vessel from ashore or on board. The Research Boat Zuse serves as a test carrier for highly automated software solutions. The Research Boat is able to the development of autonomous vessel technologies. The core of the research boat is an extended NaviBox, which can communicate with the testbed infrastructure and receive surrounded/environmental data.

3.2. Interconnected eMIR Components for Supporting the Highly Automated CPSoS Development Process

The CPS testbed interoperability architecture of eMIR is illustrated in Figure 2. This architecture provides a sensor and communication infrastructure with human-machine interaction components and enables the testing of models, implementation and physical prototypes in the loop. Closed-loop methods are particularly suitable for several reactive components, which, as an overall system, must meet complex and safety-critical requirements and test within different scenarios [10]. In order to reflect the concept of interoperability, eMIR based on the uniform data exchange format S-100 as connection between all existing elements and components. The backbone of the eMIR testbed infrastructure uses the evolving S-100 standard implemented by the uniform (canonical) data model. With S-100, the eMIR testbed can provide a sustainable connectivity and interoperability to novel systems and all compliant (prototype)

systems ensured. However, systems with common interfaces such as NMEA 0183 can also connected to the testbed via a Polymorphic Interface [11] if mainly the geographic features of the testbed have to take into account, since NMEA 0183 sentences cannot generally translated to the S-100 based reference data model. A basic idea for a generic maritime testbed is an open and adaptable design for various kinds of prototypes. Therefore, the polymorphic interface provides a compatible interface for the system under test. The interface offers the ability to integrate prototypes of various technological implementation into the infrastructure of the testbed and is highly flexible and adaptable by supporting various maritime standards, formats and regulations, such as the inter VTS exchange format (IVEF) for a vessel traffic service (VTS) system, NMEA or S-100. In order to ensure the extensibility, interoperability and flexibility of the testbed, both the virtual and the physical testbed use a middleware, which enables a standardized data exchange and where the existing components can connect and disconnect during a test run without affecting it. For the realization of different test setups, synchronous and asynchronous communication is considered. Furthermore, the middleware of the physical and virtual world can combine to perform mixed reality tests. To integrate simulative components to the physical world and vice versa, the testbed architecture proposes a simulation adapter. The simulation adapter translates the different communication protocols of the virtual and the physical testbed. The modular and extensible design of the adapter makes it possible to create an interoperable data exchange between the virtual and physical testbed. The architecture includes a static data stream processing chain consisting of the communication handler, syntax handler and semantic handler to realize the communication between virtual and physical level.

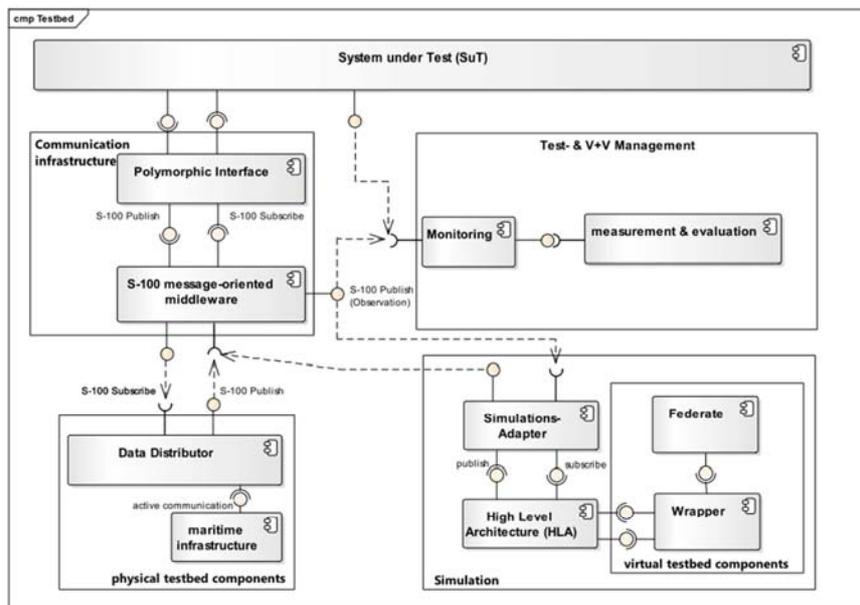


Figure 2. Testbed architecture for cyber-physical system under test and V&V management with physical and virtual testbed components.

4. Scenario Identification

An analysis of the existing requirements of the SUT to be tested became necessary in order to define scenarios in a scenario-based approach. This took place against the background of streamlining the test room in order to reduce the test effort with the help of selected scenarios. The aim was to define only relevant scenarios instead of all possible scenarios, in order to determine the quality of the system to be tested and especially its behavior in critical situations. In order to consider both the existing requirements (with their possible gaps and inaccuracies) and their criticality in the selection of relevant scenarios, a Hazard and Operability Study (HAZOP) is performed as a basis for the selection of relevant scenarios. HAZOP is a systematic search for hazards during system design (functional decomposition) and results in a criticality assessment (failure modes, effects and criticality analysis). The process ends with an assessment of the operational risk (fault tree analysis) and the derivation of relevant scenarios. We have chosen HAZOP for our work, since the architecture and complexity of components are taken into account for verification of a system. Furthermore, FMEA promotes proposals for the structure of a hardware and software system and generates preventive measures during development and operation [12]. In addition, a list of potential faults is managed by introducing a hierarchical structure in the FMEA. For each identified reason for deviation, at least one scenario is created to decide whether the system under test remains in control or not. In the event of failure or loss of control, the scenario is intended to determine the severity of the effects of the deviation (Figure 3).



Figure 3. Scenario identification.

The aim of these test methods is identification of potential risks of system components and their impact on the functionality of the system (IEC 2018) [13].

5. Use Case: A Collision Regulation Compliance Checker as System Under Test

5.1. System Under Test

In our research, we focused on the validation and verification of a ColRegChecker, which is one component of a larger system that was developed earlier, the MTCAS system

(Figure 4). MTCAS stands for Maritime Traffic Alert and Collision Avoidance System. MTCAS assists in collision avoidance by warning the ship crew while critical situations develop and recommends evasive maneuver for conflicting ships. In difference to the previous CPA- (Closest Point of Approach) calculations where a linear motion vector is calculated for each target ship based on the current speed and course, MTCAS offers functionalities for pro-active collision avoidance including methods for intelligent ship behavior prediction and an approach for cooperative collision avoidance. Different to the aircraft Traffic Collision Avoidance System (TCAS), MTCAS does not automatically intervene in terms of issuing steering commands, such that it can be seamlessly integrated into nowadays (legally regulated) operations on-board of a ship. The overall approach of this system can be found in several publications, including Denker and Baldauf [14]) and Denker, and Hahn [15].

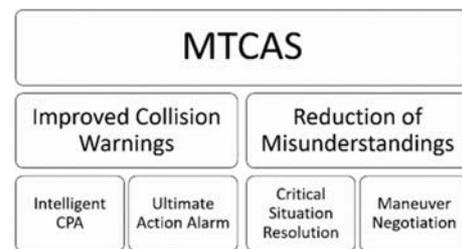


Figure 4. MTCAS principles and functionalities [16].

The upcoming subsection details this system and motivates the verification and validation of the ColRegChecker based on an expert-based systematical derivation of verification and validation scenarios based on the HAZOP method (subsec. 5.2). As the requirements for the SUT are only rudimentarily defined, they are mapped to a general requirement level based on Burmeister [17]. This specification of requirements is acceptable as it contains the restrictions on lighting according to the COLREGS (Convention on International Regulations for the Prevention of Collisions at Sea). Figure 5 shows these limitations for encounter situations of two ships: The action modes are divided into six regions with the designations A to F:

1. Assuming a heading for own vessel of 000° , own vessel is a give-way ship relative to any crossing ship in region A (005° - 067.5°) and shall alter course to starboard and avoid collision.
2. Own vessel is a stand-on ship relative to any crossing ship in region E (247.5° - 355°) and is usually not required to take any action to avoid collision.
3. If own ship is in an overtaking situation being passed by any ship from the C (112.5° - 210°) or D (210° - 247.5°) regions, it is usually required to keep course and speed.
4. Own vessel is a give-way ship relative to any ship from region B (067.5° - 112.5°) and is usually required to take action to avoid collision.
5. A head-on situation is created when own ship encounters another ship in region F (005° - 355°) and in

this situation both ships shall alter course to starboard so that each ship shall pass on the port side of the other.

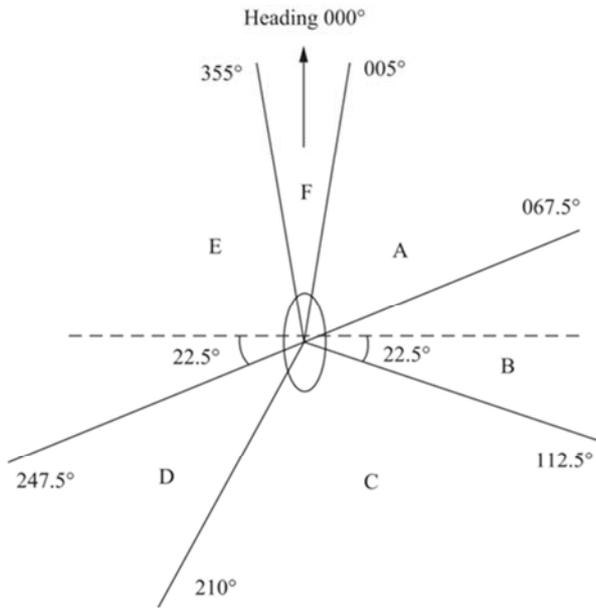


Figure 5. Overall view of basic COLREG examination rule 13 to 17 based on Burmeister [17].

In the requirements analysis the determining parameters as well as corresponding limit values were determined. So that not all values for example in 0.1 degree-steps have to be tested. The challenge was to combine similar values so that

only those that are in a certain distance from each other are tested. This discretization is realized in a special way: On one hand, all values, which (at least) occur in reality, should belong to exactly one discretization interval. On the other hand, these discretization intervals shall be smaller near the limit values, because these regions are examined more exactly.

5.2. Scenario Finding

Based on the possible reasons for deviations from the COLREGs identified in the HAZOP analysis and the angles of the right-of-way rules defined in Figure 5, 200 real-life normal traffic situations, close-range situations and ship collisions were analyzed and evaluated in order to provide a realistic reference for the scenarios to be created. In order to reduce the number of scenarios to a feasible level, the examined traffic situations are clustered and abstracted in order to clarify core situations of the rules of right of way at sea. The main commonalities in the respective situations were identified and summarized with classical methods of the black box test (equivalence classes, limit value analysis, error guessing) with regard to the possible reasons for deviations from COLREGs and right-of-way rules identified in the HAZOP analysis. The possible scenarios identified in this way were subsequently processed and calculated using navigational methods to ensure feasibility within the virtual and physical test field (Figure 6).

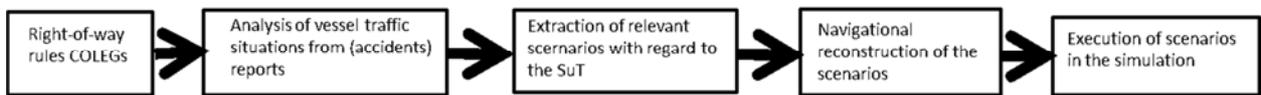


Figure 6. Schematic derivation of scenarios within the framework of a HAZOP analysis.

Table 1 below summarizes the 13 scenarios developed and the results achieved.

Table 1. Scenario selection with results.

N°	Short description	COLREG rule with bearing assignment own ship to target ship	Hybrid	Valid. Result
1a	Head on situation	Rule 14: Target ship passes with ($\alpha \geq 355^\circ$ or $\alpha \leq 5^\circ$)		$\alpha \geq 350^\circ$ or $\alpha \leq 10^\circ$ Deviation from reference: 5°
1b	Crossing situation Starboard 5°	Rule 15: Target ship passes with constant course from 5 degrees bearing from own ship through its keel line		$\alpha = 10^\circ$ limit value observation, deviation from reference: 5°
1c	Crossing situation Stb $67,5^\circ$	Rule 15: Target ship turns with course $337,5^\circ$ to course $07,3^\circ$ into the keel line of the own ship		limit value observation, limits crossing starboard: 15° - 135°
1d	Crossing situation Stb $112,5^\circ$	Rule 15: Target ship moves from bearing $202,5$ with course $22,5^\circ$ to $43,7^\circ$ towards Keel line own ship		limit value observation, limits crossing starboard: 15° - 135°
1e	Overtaking situation 210°	Rule 13: Target ship moves at double speed of own ship from bearing 210° from own ship on course 015°		limit value observation, deviation from reference, limits are: $157,5^\circ$ - $202,5^\circ$
1f	Overtaking situation $247,5^\circ$	Rule 13: Target ship moves at double speed of own ship from bearing $247,5^\circ$ from own ship on course $067,5^\circ$		limit value observation, Deviation from reference, limits are: $157,5^\circ$ - $202,5^\circ$
2	Overtaking or crossing (1) situation	Rule 13, 15, 16, 17: Different assessment of the situation. Wind from southeast with 25 knots in rough seas		Due to the influence of wind, the ColRegChecker output constantly changing assignments between crossing, overtaking and no danger (see Figure 12)
3	Collision situation	Communication by radiotelephony. Target vessel is attempting to pass the own ship in front of their bow		Successful reconstruction of a mutually contradictory situation assessment by two ships
4	CPA calculation	Changes in course of the target ship are not immediately detected		The CPA/TCPA calculation does not always meet the current requirements of safe maritime traffic
5	Weak and wrong reactions	Rule 2: Both ships turn with their courses into each other		Successful reconstruction of weak and wrong reactions

N°	Short description	COLREG rule with bearing assignment own ship to target ship	Hybrid	Valid. Result
6	Overtaking situation	Rule 13: Scenario with 4 ships	X	Successful enrichment of the real test field with 2 real and 2 virtual ships. ColRegChecker processes only 2 ships
7	Traffic Separation Scheme (TSS)	Rule 10: Recognition of a traffic separation area		TSS is not recognized by ColRegChecker
8	Fishing vessels in a TSS	Rule 8, 10, 15, 16, 17: Scenario with 5 ships in a TSS	X	Successful enrichment of the real test field with 2 real and 3 virtual ships. ColRegChecker processes only 2 ships
9	In sight (Overtaking or crossing (2))	Different evaluations of the situation by own and target ship		The "in sight" condition is not recognized by ColRegChecker
10	Narrow Channel	Rule 9: Narrow channel crossing		Narrow Channel is not recognized by ColRegChecker
11	Safe speed	Rule 6, 8, 19: Scenario with 5 ships	X	Successful enrichment of the real test field with real and 3 virtual ships. ColRegChecker processes only 2 ships. Reduced visibility is not recognized by ColRegChecker.
12	Reduced visibility	Rule 3, 6, 7, 8, 19: Scenario with 2 ships in reduced visibility		The own ship in the test field did not have a sufficient rotating speed to carry out the full scenario
13	Multi-Ship Scenario	Scenario with 6 ships in a complex traffic situation.	X	Successful enrichment of the real test field with 5 virtual ships. ColRegChecker processes only 2 ships

5.3. Simulator Based Testing

The TRANSAS NTPro simulator software is used to simulate the 13 scenarios developed. This software provides among others the NMEA data sentences VTG (Track made good and Ground speed), GGA (Global Positioning System Fix Data, Time, Position and fix GPS receiver) and RMC (Recommended Minimum Navigation Information). The "RMC" sentence is used to extract the position (latitude and longitude), speed and heading. Additionally, the position data are converted from degree minutes to decimal degrees.

The NMEA data are available in the raw data layer (Figure 7) in their original form and unchanged to make them available to other systems, for example. Via the data model,

the transformation to S-100 (S-100 Universal Hydrographic Data Model ISO/TC211) takes place. To use this data as input for the SuT (ColRegChecker) a re-transformation into NMEA data must be performed. Using the NAS data recording, a comparison of the NMEA data can be made for falsification and rounding by the S-100 transformation. The manipulation-software is used within the simulation to increase the realism of the scenarios. This is done for example by noisy linear courses or jumps in AIS positions. Within the physical test field, the manipulation- SW allows data streams to be generated by data manipulation or data generation that deviate from the real scenario. Thus, for example, virtual ships or error injections are import for a test to check the robustness of the SuT.

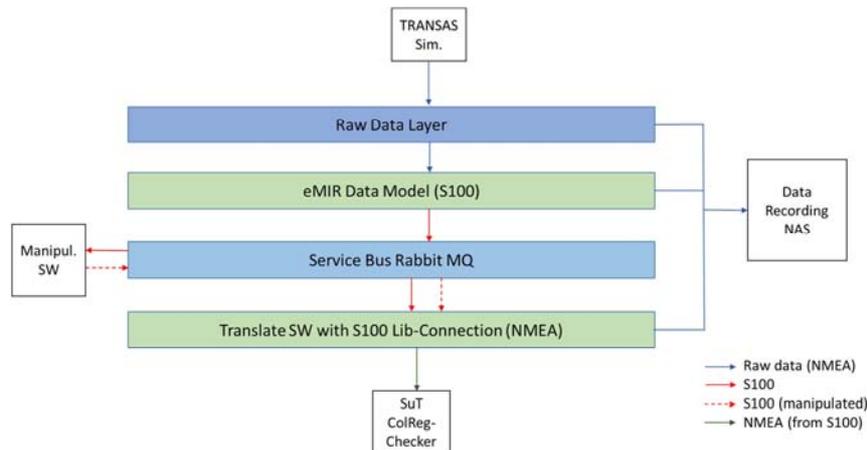


Figure 7. eMIR architecture for simulation.

5.4. Setting up and Carrying out the Physical Tests

The scenarios from Table 1 were each carried out in the simulation using the eMIR architecture and subsequently performed in the test field with max. 2 real ships. In the scenarios in which more than two ships were required, the real situation was enriched with virtual ships (Manipulation Sw., Figure 7). To realize the scenarios within the test field, all sensors required for testing the SuT were used. Therefore eMIR

provides the sensor information of the connected sensors Radar, AIS, wind sensor, a DGPS (Differential Global Positioning System) for positioning, weather chart recorder (Navtex) and Inertial Measurement Unit (IMU). All Sensor data were provided by NMEA 0183 or 2000. In September 2019, the mobile eMIR container was mounted on the research vessel DENEBO of the BSH (Federal Maritime and Hydrographic Agency) to carry out the scenarios in the test field (Figure 8).



Figure 8. Mobile eMIR Platform on deck DENEb.

The sensors used (Figure 9) were installed on the mobile eMIR container and on deck DENEb. The DGPS was installed on the container roof to determine the position of the own ship. For tracking the target ship the NaviBox with radar, AIS, and a camera was installed on the bow of the DENEb. A wind sensor was installed there in addition. The acquired sensor data were sent to the container via LAN, recorded there and subsequently evaluated. The LIDAR Riegl VZ-2000 was installed on the port side on a stabilized platform to detect close-range situations. The drone with its camera was launched from the aft deck and served to observe and record the scenarios from the sky. The Drone has an internal recording and was not integrated into the physical test field.

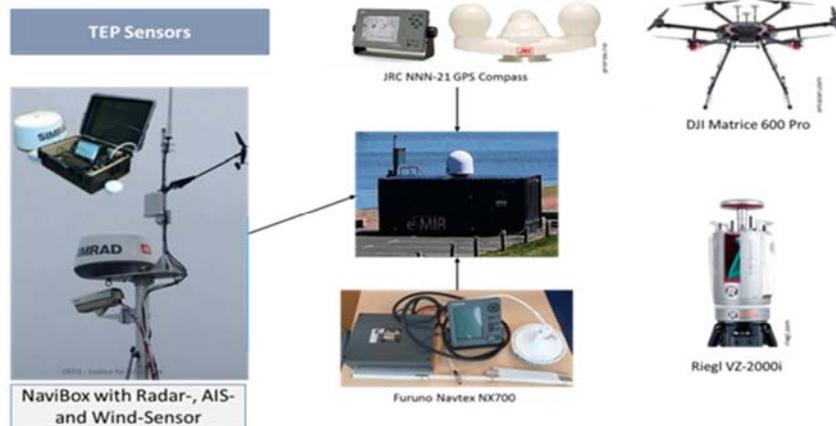


Figure 9. Sensors installed on board.

6. Evaluation and Results

6.1. Simulation Results

In the execution of the defined scenarios (Table 1), there were significant differences between the manually navigational calculated course and speed data of the participating ships and those in the simulation, as these ships were assumed as point targets within the manual calculation, but ship models with defined behavior were used in the simulation. Although the ship models used were similar to the ships to be used in the real test, ideally simulation models of ships should be used which correspond to the behavior of ship class of the ships in the test field, so that there are no deviations which have to be calculated out afterwards. The recognition of the existing traffic situation by the SuT (heading, crossing or overtaking) took place at different times between the navigational calculation and the simulation, but did not influence the basic recognition by the SuT. Within the simulation, the limits of the SuT were tested using equivalence classes and limit value considerations within the scenarios. The summary of the overall result of the simulation test is shown in Figure 10. The ColRegChecker assigns an encounter situation between the own ship (OS) and any target ship (TS) to one of the following situation type labels: "Head-On", "Overtaking", "Being-Overtaken", "Crossing", "No-Danger". A "Head-On" situation exists if the

TS is in front of the OS in the direction of the course and the course difference of OS and TS is between 170° and 190° . In a "Head-On" situation, the TS is in front of the OS in course direction, provide that the OS course does not deviate more than 10° from the bearing to the TS.

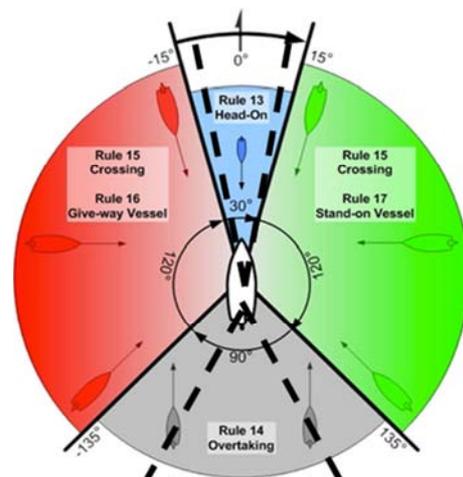


Figure 10. Simulation overall result.

An "overtaking" situation exists if the TS is in front of the OS in course direction and the speed of the OS is higher than that of the TS. An "Overtaking" situation exists, if the TS is in front of the OS in course direction, if the course difference

is smaller than 22.5° and larger than 337.5° and the OS course does not deviate more than 45° from the bearing to the TS. If the OS speed in this case, however, is lower than the TS speed, a "No-Danger" situation exists. A "leg-overtaken" situation exists, if the TS is behind the OS in course direction and the speed of the OS is lower than the speed of the TS. In a "Legg-Overtaking" situation the TS is behind the OS in course direction, if the course difference is smaller than 22.5° and larger than 337.5° . However, if the OS speed is greater than the TS speed in this case, a "No Danger" situation exists. A "crossing" situation exists if the above conditions are not fulfilled. As an example of the simulation results, the Figure 11 shows the results of scenario 1a. In this figure as a resulting example from the ColRegChecker the tracks of the own ship (OS) are marked with blue color and those of the target ship (TS) with red color.

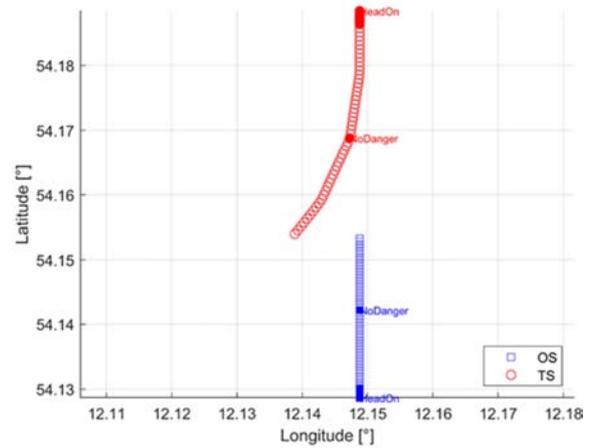


Figure 11. Result simulation scenario 1a.

Probably due to the simulated wind influence, the assignment of the ColRegChecker constantly changes back and forth between crossing, no danger and overtaking (Figure 12). Further analyses on this are in progress.

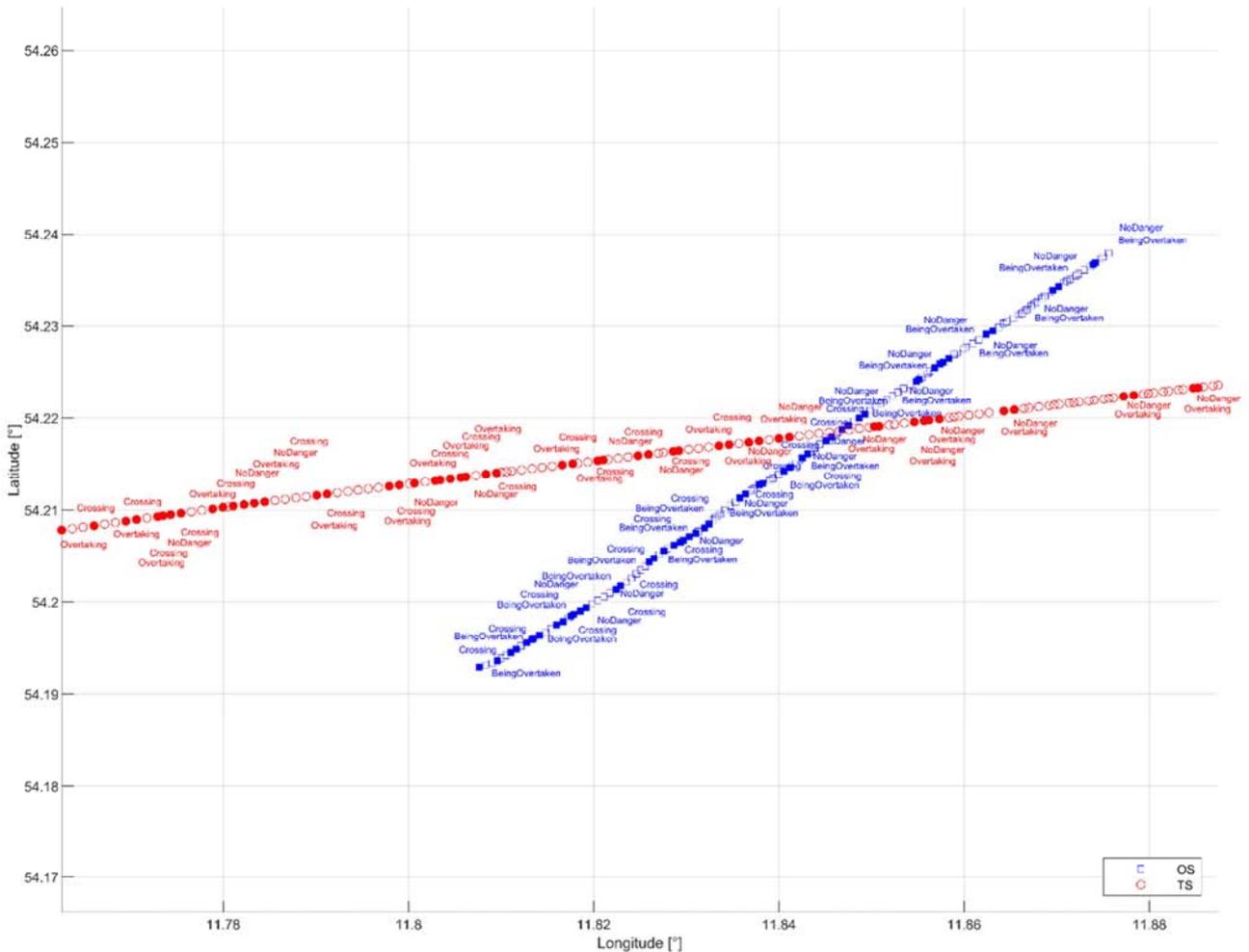


Figure 12. Result simulation scenario 2 (constantly changing assignment by the SuT).

In contrast, in scenarios without environmental influences, for example in a "Head On" situation (Figure 11), a clear assignment is made and displayed for both ships. After the

situation has been clarified by the target ship, "No danger" is displayed for both ships. According to COLREGs, the own ship would have had to move to starboard in this situation as

well. Due to the data storage on the NAS (Network Attached Storage) (Figure 7) it was possible to compare the sensor NMEA raw data with the eMIR data to determine that there are no differences between these data.

6.2. Results in the Physical Test Field

Of the 13 developed scenarios, 12 could be carried out. It was not possible to carry out scenario 12 because the DENEb (own ship) could not reach the required rate of turn. Due to the extremely calm weather conditions, the influences of the environment on the classification of the right-of-way could not be tested as a result of the ColRegChecker. All in all, the execution of the scenarios in the Baltic Sea test field

revealed significant differences between the results of the simulation and those in the test field, as the associated ship models in the simulation did not correspond to the dynamic behavior of the DENEb and the target ship. In these cases, identical results can only be achieved if the ship models used in the simulation correspond to those in reality. The added value of the manipulation SW was clearly demonstrated in the test field. Thus, the multi-ship scenario (scenario 13) in Figure 13 could be performed with only one real ship (own ship, black vector) and 5 different virtual ships (colored vectors) in the test field. The start conditions of this scenario are shown in bottom side of the image and are visualized on the upper side.

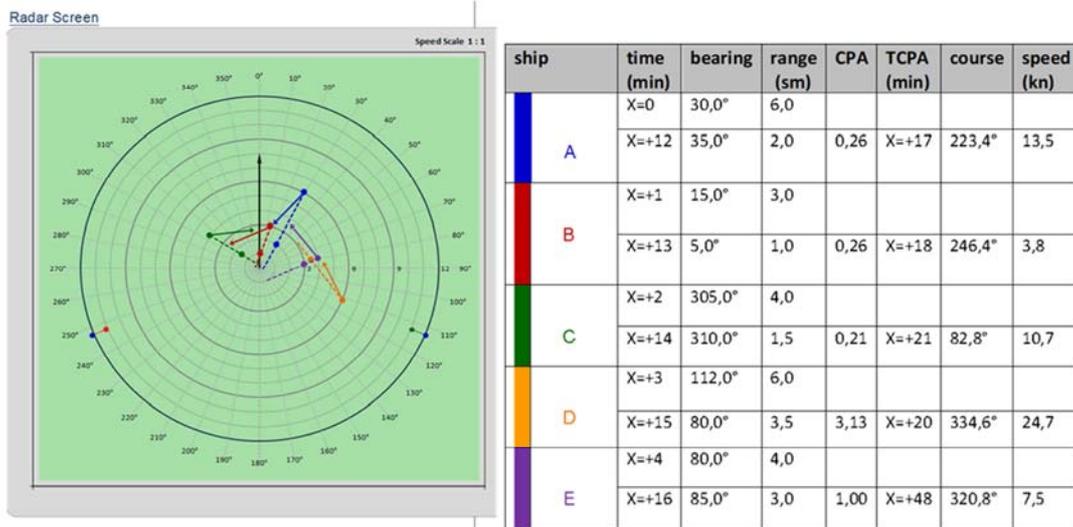


Figure 13. Start conditions scenario 13.

During the run of the scenario, none of the ships A-E (virtual ships) makes a course or speed change to avoid a close-range situation, so that the own ship has to take the initiative. The required evasive maneuver of the own ship is

carried out at time X=+31min. At this time, the own vessel goes on a course of 330° and increases its speed to 12kn. The results resulting from this maneuver is shown in Figure 14.

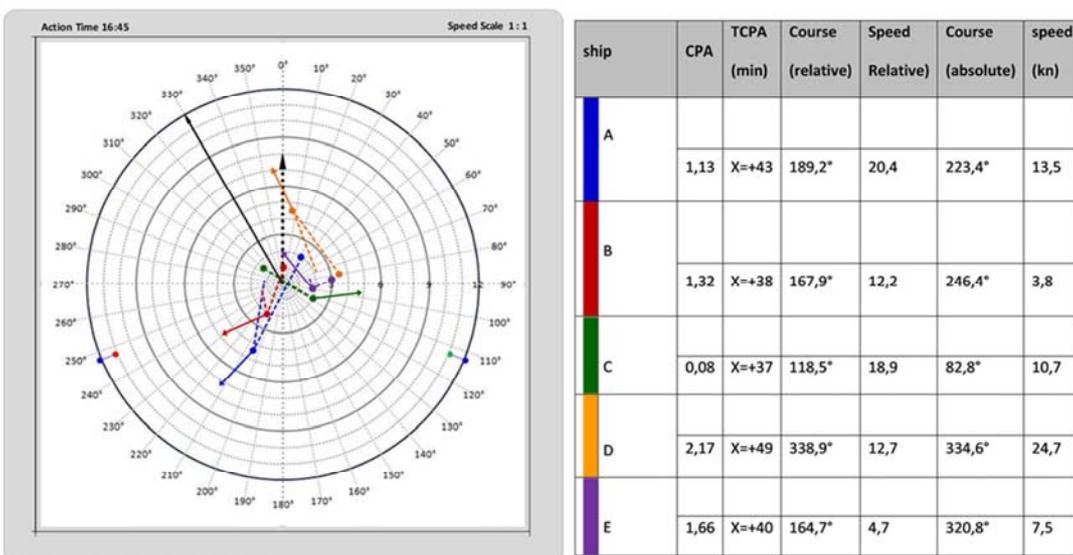


Figure 14. Changing course and speed of the own ship to to avoid a close-range situation.

The ColRegChecker label ("Head-On", "Overtaking", "Being-Overtaken", "Crossing", "No-Danger") are only assigned if the distance between the vessels involved in the Closest Point of Approach (CPA) is less than two nautical miles.

6.3. Results and Critical Discussion

It was found that the developed architecture takes into account the V&V requirements of different application systems and at the same time enables the verification and validation of complex systems or SoS. The use of the same data model in simulation and physical experiments allowed a seamless transition from fully virtual tests to physical tests. The verification of the specific properties of SoS both in the simulation and in the field as well as in the combination of both could be performed effectively. All test field components, both virtual and physical, were able to process the same data and, in particular, make it available for a wide range of applications. The use of the polymorphic data interface, which enables the provision of sensor data, situation pictures etc. in a form suitable for the SoS under test, contributed to this in particular. With regard to the execution of scenario-based tests, the fundamental question arises whether the scenarios performed in the virtual test field must be repeated in their completeness in the physical test field. Although we have done this in 2019 in order to obtain experience and opportunities for comparison, our answer to this question is clearly: no. Especially for systems with emergent system behavior, the number of scenarios required to achieve acceptable test coverage can be very high. This coverage cannot be achieved in a physical test field, especially in the maritime domain, due to resource constraints. Thus, the number of scenarios to be performed in the physical test field must be limited to a number of particularly critical or in result uncertain scenarios. The findings from the real environment can in turn be used to optimize the test methods and procedures in the simulation and on the test benches. In the long term, this could be linked to the goal of optimizing the test procedures to such an extent that the approval (homologation) of complex systems (vessel bridge system, assistance systems) is largely possible without the need for costly real tests. In our case of the complete repetition of the scenarios in the test field, this caused additional questions and an increased need for analysis, since the dynamic ship behavior of the own ship and that of the target ship did not correspond to the models available in the simulator. Thus bearings, distances between the two ships were delayed or reached only after some corrections. With regard to the manual creation of the scenarios and their mapping to the requirements of the SuT, it should be noted that this was extremely time-consuming and is not feasible in this form for more complex systems. However, the navigational calculation as a basis for scenario execution has proven to be useful and profitable for proving concrete situation representations and process controls within the simulation. The balance of virtual and physical testing is a

key concern in reducing design time and cost. Integrating physical and virtual testing is more than process optimization of time and cost. It contributes to recasting the design process in response to changes in customer requirements as well as to design changes which arise during testing. The importance of AR (Augmented Reality) for scenario definition within the virtual and physical test field should also be underlined. In the field of validation and verification, mixed reality is able to significantly increase the efficiency and coverage of tests without losing touch with reality. In contrast to virtual reality, augmented reality does not create a new world parallel to reality, but rather expands physical reality through techniques. AR places an interactive, virtual layer on the environment surrounding us, thus providing additional information to the existing real perceptions. Especially in the maritime domain the AR plays an important role, because by means of the transfer of digital (virtual) data into reality for example safety-critical close-range situations of several ships can be carried out, which could not be carried out purely realistically on the one hand due to safety aspects and on the other hand due to cost reasons. Regarding the COLREG checker it was found that:

1. the important criterion "in sight" was not detected (COLREGs rules 4-10)
2. special sea areas such as "narrow fairways" or TSS (Traffic Separation Scheme) are also not detected (COLREG rule 9 and 10)
3. the speed of vessels involved has no influence (keyword "safe speed") (COLREG rule 6)
4. Restricted visibility was not recognized (Rule 19 COLREGs)
5. Exemptions are not identified (Rule 38 COLREGs).

It was not tested whether these conditions are implemented within other components of MTCAS. The influences of environmental conditions (wind, currents, etc.) on the system behavior of the SuT could not be tested in the physical test field, as these parameters were more or less non-existent due to an absolutely calm sea.

7. Conclusion

Up to now, system behavior in maritime transport has been regarded as a stochastic process. This corresponds to the attempt to cover the state domain in a representative way by pure sailing. From the experience and knowledge gained so far, a turnaround in method development towards simulation-based approaches for determining and verifying functional limits is necessary. The testing of complex systems or SoS requires a change from the manual functional based scenario generation (we develop in 2019) to an automated scenario approach using a selection of historical stock data. Due to the high level of parameterization of systems and their diverse dependencies, the number of scenarios to be created will also increase strongly, so that the achievement of acceptable test coverage can only be achieved in an economically justifiable manner by means of automated scenario generation. The

physical test field supports the verification and validation (V+V) during the whole system development process and enables the testing of a system under test in a simulative and/or physical test environment. The heterogeneity of the systems to be tested and their extensive requirements due to the system development methodology for these novel systems pose challenges for the design of current maritime test setups. In order to enable system developers and test engineers to efficiently test the automated ship guidance systems, the implemented test field provides an answer to the question "How must a physical test field be designed to efficiently support verification and validation in the system development of automated ship guidance systems? The combination of the virtual and physical test field is able to significantly increase the test coverage.

Acknowledgements

Part of the research is based on the outcome of the project ACTRESS which is supported by the German Federal Ministry of Economic Affairs and Energy.

References

- [1] T. Porathe, "ACCSEAS Baseline and Priorities Report," ACCSEAS Project, 2013.
- [2] T. Porathe, M. Lützhöft, and G. Praetorius, "Communicating intended routes in ECDIS: Evaluating technological change," *Accid. Anal. Prev.*, vol. 60, Nov. 2013, pp. 366–370.
- [3] DMA, "An overview of the 'Maritime Cloud' – proposed information exchange infrastructure for e-navigation". Copenhagen 2013, SNPWG17-9.3.
- [4] MONA LISA, "Dynamic and Proactive Routes or 'green-Routes'", Sjöfartsverket Sweden, Stockholm.
- [5] G. Mannarini, G. Coppini, P. Oddo, and N. Pinardi, "A Prototype of Ship Routing Decision Support System for an Operational Oceanographic Service," *TransNav Int. J. Mar. Navig. Saf. Sea Transp.*, vol. 7, no. 2, 2013, pp. 53–59.
- [6] K. Aichhorn, de la Cuesta de Bedoya, P. Berglez, M. Lopez, M. Troger, and A. Kemetinger, "Maritime Volumetric Navigation System," in *Proceedings of the 25th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS 2012)*, Nashville, Tennessee, 2012.
- [7] SESAME Strait. <http://straits-stms.com/Project.html>.
- [8] H. Adono, "Smart Ship Application Platform Project (SSAP Project)". [Online]. Available: <http://www.mlit.go.jp/common/001039009.pdf>. [Accessed: 04-May-2020].
- [9] L. Schnieder, and R. Krenkel, "Betreibermodell einer Forschungsinfrastruktur für die Entwicklung intelligenter Mobilitätsdienste im realen Verkehrsumfeld". 16. Symposium Automatisierungssysteme, Assistenzsysteme und eingebettete Systeme für Transportmittel (AAET), 12.-13. Feb. 2015, Braunschweig. ISBN 978-3-937655-34-5.
- [10] A. Rizvanolli, H-C. Burmeister and O. John, The Role of the European Maritime Simulator Network in Assessing Dynamic Sea Traffic Management Principles TransNav, the *Int. J. on Marine Navigation and Safety of Sea Transportation* 9 559–64, 2015.
- [11] M. Brinkmann, A. Hahn, Physical Testbed for Highly Automated and Autonomous Vessels 16th Int. Conf. on Computer and IT Applications in the Maritime Industries COMPIT, 2017.
- [12] J. J. Stadler, and N. J. Seidl, "Software failure modes and effects analysis". In *Reliability and Maintainability Symposium (RAMS)*, 2013 *Proceedings-Annual* (pp. 1–5): IEEE.
- [13] IEC 2018 Failure modes and effects analysis (FMEA and FMECA).
- [14] C. Denker, M. Baldauf, S. Fischer, A. Hahn, R. Ziebold, E. Gehrman, and M. Semann, e-Navigation based cooperative collision avoidance at sea: The MTCAS approach. *European Navigation Conference (ENC)*, 2016.
- [15] C. Denker, and A. Hahn, "MTCAS: An e-Navigation Assistance System for Cooperative Collision Avoidance at Sea", 2017.
- [16] M. Steidel, A. Hahn, "MTCAS – An Assistance System for Collision Avoidance at Sea". *Conference Paper March 2019*, <https://www.researchgate.net/publication/332111082>.
- [17] H. Burmeister, and W. Bruhn, „Designing an autonomous collision avoidance controller respecting COLREG” in *Maritime-Port Technology and Development*, 2015.