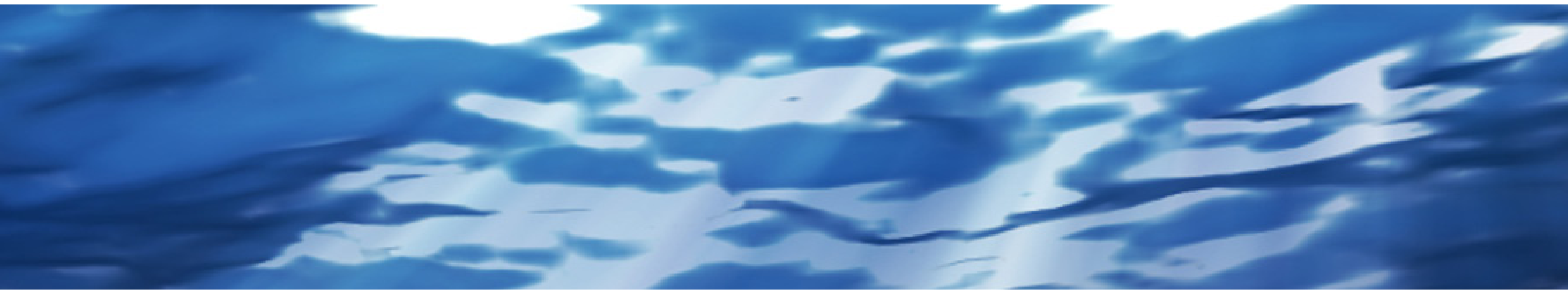
LIGHTHOUSE

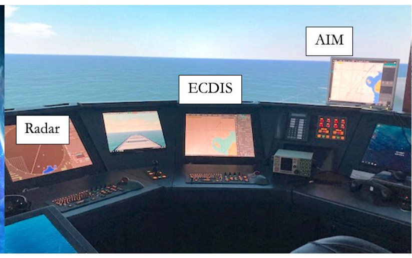
SWEDISH MARITIME COMPETENCE CENTRE



LIGHTHOUSE REPORTS

Operationalizing COLREGs in SMART ship navigation

Understanding the limitations of algorithm-based decision support systems in traffic situations



A research project carried out within the Swedish Transport Administration's industry program Sustainable Shipping, operated by Lighthouse, published in April 2022

TRAFIKVERKET

LIGHTHOUSE
SWEDISH MARITIME COMPETENCE CENTRE

HÅLLBAR SJÖFART

Operationalizing COLREGs in SMART ship navigation

Understanding the limitations of algorithm-based decision support systems in traffic situations

Authors

Reto Weber, Chalmers

Katie Aylward, Chalmers

Scott MacKinnon, Chalmers

Monica Lundh, Chalmers

Mikael Hägg, RISE

A research project carried out within the Swedish Transport Administration's industry program Sustainable Shipping, operated by Lighthouse

Summary

The maritime industry is undergoing a transformation driven by digitalization and connectivity. The technological realization of Maritime Autonomous Surface Ships (MASS) presents significant challenges for the maritime human factors research community. These challenges relate to system design, human-automation interaction, stakeholder training, use and acceptance of new technology systems, and on a larger scale, how the regulatory framework, including the Collision Regulations (COLREGs) will be impacted within a MASS system. Decision support is the next step in the transformation towards more connected ships, however, such systems for navigation are largely unexplored from the users' perspective.

The decision support system studied in this project was developed by Wärtsilä and is called Advanced Intelligent Manoeuvring (AIM), aligning with "low-level automation" or Level 1 (out of a 4-level progression) of MASS. AIM can generate suggestions for course or speed alterations based on data from surrounding traffic. A full-mission bridge simulator study was conducted at Chalmers University of Technology in Gothenburg, Sweden with nineteen Swedish navigators. Three traffic scenarios each with three ships were completed in both baseline (no AIM) and AIM conditions. A mixed methods data collection and analysis approach was employed using questionnaires, collective interviews, and an evaluation of the ship tracks.

The results show that the navigators perceive AIM as an advisory tool, to visualize how traffic situations could unfold, an outcome currently difficult for most navigators to conceive. This report discusses the present and near future of the maritime sociotechnical system, highlighting the benefits of automation, while remaining vigilant about the potential dangers.

Sammanfattning

Den maritima industrin genomgår en transformation som drivs av digitalisering och uppkoppling. Det tekniska förverkligandet av Maritime Autonomous Surface Ships (MASS) innebär betydande utmaningar för forskarsamhället inom Maritime Human Factors. Dessa utmaningar relaterar till systemdesign, interaktion mellan människa och automatisering, utbildning, användning och acceptans av nya tekniksystem, och i större skala, hur regelverk, inklusive kollisionsreglerna (COLREG) kommer att påverkas inom ett MASS-system. Beslutsstöd är nästa steg i transformationen mot mer uppkopplade fartyg, men sådana system för navigering är till stor del outforskade ur användarnas perspektiv.

Beslutsstödssystemet som studerades i detta projekt har utvecklats av Wärtsilä och kallas Advanced Intelligent Maneuvering (AIM), i linje med "lågnivåautomation" eller nivå 1 (av en 4-nivå progression) av MASS. AIM kan generera förslag på kurs- eller hastighetsändringar baserat på data från omgivande trafik. En full-mission bryggsimulatorstudie genomfördes vid Chalmers tekniska högskola i Göteborg med nitton svenska navigatörer. Tre trafikscenarier vardera med tre fartyg genomfördes i både baslinje (ingen AIM) och AIM-förhållanden. En blandad metod för datainsamling och analys användes med hjälp av frågeformulär, kollektiva intervjuer och en utvärdering av fartygets spår.

Resultaten visar att navigatörerna uppfattar AIM som ett rådgivande verktyg för att visualisera hur trafiksituationer kan utvecklas vilket för närvarande är svårt för de flesta navigatörer att föreställa sig. Denna rapport diskuterar det maritima sociotekniska systemets nuvarande och nära framtid, och lyfter fram fördelarna med automatisering, samtidigt som man är vaksam på de potentiella farorna.

Contents

Table of Figures

Tables

List of abbreviations

AIM	Advanced Intelligent Manoeuvring
AIS	Automatic Identification System
ARPA	Automatic Radar Plotting Aid
COLREG	International Regulations for Preventing Collisions at Sea, 1972
CPA	Closest Point of Approach
ECDIS	Electronic Chart Display and Information System
FMBS	Full Mission Bridge Simulator
NTPro	Wärtsilä's bridge simulator software Navi Trainer
OS	Own ship
OOW	Officer of the Watch
SA	Situation Awareness
TCPA	Time to Closest Point of Approach
TG	Target ship
VHF	Very High Frequency radio communication
VTS	Vessel Traffic Services
WS	Workstation
XTD	Cross Track Distance

1 Introduction

Smart vessels instrumented with automated systems requires a reliance upon algorithm-based solutions that infer interpretations of the COLREGs given a specific traffic situation. In reality, decisions made on manned ships are influenced by the experience of the operator and his/her interpretation of good seamanship. Furthermore, the geographical area, traffic pattern, complexity of the traffic situation, ship type, as examples, will influence operator decision-making and application of COLREGs. As a step towards developing partly or fully autonomous ships, various companies are providing ship owners with software which analyses the traffic situation and recommends or "suggests" to the officer of the watch an action to avoid a close quarters situation. In this project, Wärtsilä's AIM (Advanced Intelligent Manoeuvring) application was integrated in Chalmers Full Mission Bridge Simulator (FMBS) with the purpose to:

- Understand the effect of an algorithm-based decision support system (AIM) on operator decision making and overall navigation practices.
- Contrast and associate the plan and decision made by a human operator with the suggestion provided by AIM specifically related to COLREG Rule 8, which states that any action taken to avoid collisions shall, if the circumstances of the case admit, be positive, made in ample time and with due regard to the observance of good seamanship.
- Contrast and associate the expectations on other ships' (predicted behaviour) to solve a traffic situation with the suggestion(s) provided by AIM on the other ships.
- Understand the level of agreement and potential conflicts of decisions made and actions taken between rule-based machine algorithms and human applications of these rules in traffic situations.

The overall goal of the project was to provide answers to questions such as:

- Can experience, good seamanship, situational awareness and other non-technical skills that impact upon safe and efficient navigation be implemented into algorithms?
- Can traffic situations that have *both* human operators and smart/autonomous vessels be safely resolved compared to traditionally manned navigation bridges?
- From the navigator's perspective, what is the role of a collision avoidance decision support system for navigation?
- From the navigator's perspective, what are the benefits and potential challenges associated with a decision support collision avoidance system?

In essence, can an algorithm reflect the meaning and implication of COLREG Rule 8?

2 Background

There exists a conflict between the push towards autonomy and the need for it in the maritime industry. This may sound controversial depending on the background and role of the reader in relation to the maritime industry. In 2018, the Nautilus Federation collected data from almost 1,000 maritime professionals across a dozen countries asking about critical issues surrounding the adoption of autonomous vessels. The results revealed that 84% of maritime professionals consider automation to be a threat to seafaring jobs, and 85% believe that unmanned remotely controlled ships pose a threat to safety at sea (Nautilus 2018). However, certain stakeholders within the maritime industry consider themselves to be ready for innovation and the adoption of more autonomous systems. Within the Nautilus survey, more than 60% of respondents felt that automation has the potential to make certain aspects of the shipping industry safer if introduced systematically, slowly and with a priority focus on the human-automation relationship (Nautilus 2018).

It is important to acknowledge the technological developments in the past several decades, particularly in the domain of artificial intelligence (AI) and machine learning. These technologies can and will change the shipping industry. The digitalization of certain aspects of the maritime industry have already proved beneficial, particularly from shore side operations including, port and container handling activities (Michaelides, Lind et al. 2020). Moreover, there are several countries hosting national initiatives for autonomous ships, including live test beds in Norway, China, Finland, and Germany. Norway is the most famous as the first test bed and remains the leader in autonomous ship development with three designated autonomous ship test beds across the country, one of which hosts the famous Yara Birkeland.

Today, most funding agencies and research initiatives prioritize projects which hypothesize a future with a remote-control centre (RCC), similar to the MUNIN concept (Man 2015), or other advances to greater levels of autonomy (Ramos, Utne et al. 2018, Ramos, Thieme et al. 2020). There is an underlying assumption in most existing research exploring remote and autonomous ship cases that the challenging questions of how to incorporate seafarer experience, and seamanship into artificial intelligence will be solved (Abilio Ramos, Utne et al. 2019, Ramos, Thieme et al. 2020, Porathe 2021). Planning for and envisioning possible realities outside the scope of existing regulations is important, as this leads to innovation. However, these challenging questions, involving the present and foreseeable future, are crucial to safe automation integration at any level while humans remain a part of the system.

3 Existing Research

For the last decade there has been substantial research efforts to develop collision avoidance algorithms, motivated by the goal of autonomous or remote-controlled ships (Woerner 2016, Woerner, Benjamin et al. 2016, Zhang and Furusho 2016, Abilio Ramos, Utne et al. 2019, Perera and Batalden 2019, Woerner, Benjamin et al. 2019, Ramos, Thieme et al. 2020). These works have attempted to quantitatively evaluate and implement the subjective nature of the COLREGs (IMO 1972) through various approaches including optimization methods, reinforcement learning, fuzzy-logic, neural networks, and Bayesian networks (Woerner, Benjamin et al. 2019, Porres, Azimi et al. 2021). As machine learning and more advanced neural networks are developed, the potential for collision avoidance systems will be further advanced.

However, the application of AI for autonomous vessels is still in its earliest stages simply because of the complexities of navigation even in the simplest of traffic situations. A recent review paper for AI in collision avoidance systems identified that only 48% of the studies reviewed complied with the COLREGs, creating a gap between ongoing research and the requirements by regulatory frameworks (Porres, Azimi et al. 2021). AI for collision avoidance systems is not an exclusive property of unmanned ships; collision avoidance algorithms may also be used for manned ships as decision support systems. Decision support systems for navigation are the next systematic step towards smarter ships while humans remain in control of the ship.

While decision support has been well studied in other domains, there are few research efforts studying decision support systems for maritime applications from the operator perspective. Only a handful have been developed and tested on end-users over the last decade, none of which appear to be in commercial use today. One example is NAVDEC or navigation decision supporting system, which is the first of its type on the market. Originally proposed in 2012 (Pietrzykowski and Wołejsza 2016) NAVDEC plans manoeuvers for the navigator that comply with COLREGs and is based on set distances and times. Although promising, it seems that there hasn't been any updated research or movement on the NAVDEC website since 2019 (http://navdec.com/en/).

Another like system is Multi-ARPA (MARPA) which provides the navigator information on safe headings, and operates based on an algorithm designating direct hazards for the Own Ship (OS) for the set of manoeuvers in an encounter situation of numerous ships (Ożoga and Montewka 2018). This paper highlighted the potential of decision support systems for navigation based on two scenarios, which is comparable to the present study. This paper also highlighted the need to proceed with caution and the difficulties of writing accurate algorithms given the challenging operational environment of ships (Ożoga and Montewka 2018). There has been no further work about MARPA since its introduction in 2018.

The Sea Traffic Management (STM) project developed "STM-services" for use both on board and on shore which provide additional information about surrounding traffic and allow information exchange between ships. The goal was to enable operators to make more informed decisions based on real-time information. The results from the STM project indicate that although several of the STM-services were useful, they caused changes to existing work practices which could impact safety, communication structures, workload, and situational awareness (Aylward 2020, Aylward, Johannesson et al. 2020, Aylward, Weber et al. 2020). This work highlighted the need for an increased research focus on the potential impacts on the sub-systems within the maritime sociotechnical system.

The lack of relevant research is surprising given the significant stakeholder focus on higher levels of MASS i.e., Levels 2-4, which includes aspects of fully autonomous and/or remote-controlled operations. Ironically, this has created the knowledge gap of lower levels of automation needed to inform the research of higher levels of autonomy. Decision support systems should improve the safety of navigation; however, the reality is that any additional technology added onto the bridge can have unwanted or surprising consequences (Bainbridge 1983). This technology can lead to over-reliance, misunderstandings, and even conflict between the human operator and automation (Endsley 1995, Lee and See 2004, Endsley 2017).

4 Advanced Intelligent Manoeuvring (AIM)

4.1 General description

The decision support system studied in this project is called Advanced Intelligent Manoeuvring (AIM). AIM is a decision support system developed by Wärtsilä that provides suggestions for collision and grounding avoidance. To prevent collisions (or near misses) in traffic situations, navigators are bound to follow The International Regulations for Preventing Collisions at Sea (COLREGS) which are basically the "rules of the road" for ships and other vessels at sea i.e., making it clear as to which ship is the "stand on" and "give-way" ship and what correct action should be taken to avoid a collision.

To support the navigator in ascertaining if a risk of collision exists, ARPA (Automatic Radar Plotting Aid) and AIS (Automatic Identification System) are mainly used. ARPA is a radar with capability to track and obtain information about plotted targets (TG) such as (among others) the Closest Point of Approach (CPA) and the time to CPA (TCPA) and also includes a Trial Manoeuver function where the effect of an own ship manoeuver on all tracked TGs can be simulated. AIS is an automatic tracking system in which ships transmit information about the ship itself such as name, position, size, course and speed, etc. to other AIS receivers and can be depicted on both the radar and the Electronic Chart and Information System (ECDIS). AIS is regarded as useful source of supplementary information to that derived from other navigational systems (including radar) and therefore an important “tool” in enhancing situation awareness in traffic situations.

AIM is marketed as a smart addition to a standard ARPA and Trial Manoeuver covering all working cycles of operations, including situation monitoring, problem detection, suggesting a manoeuver and monitoring execution of the manoeuver based principally on mathematical calculations (Wärtsilä 2020). Based on the assumption that the other ships keep their course and speed, AIM provides graphical solution(s) on how to solve the traffic situation either by changing own ship’s course or reducing speed. The platform includes an additional feedback system that “plays ahead” the manoeuver before its execution. It needs to be noted that the software is still being further developed and that the following description is based on the available software version used during the trials. The application performs the following functions (Wärtsilä 2019):

- Producing a system analysis and informs the watch officer of situations in which a collision of ships is possible.
- Calculating a manoeuver recommending the course and/or speed for avoiding collision with dangerous targets in compliance with the COLREGs.
- Displaying manoeuvring suggestions graphically and textually, on the screen (see figures 1 and 2).

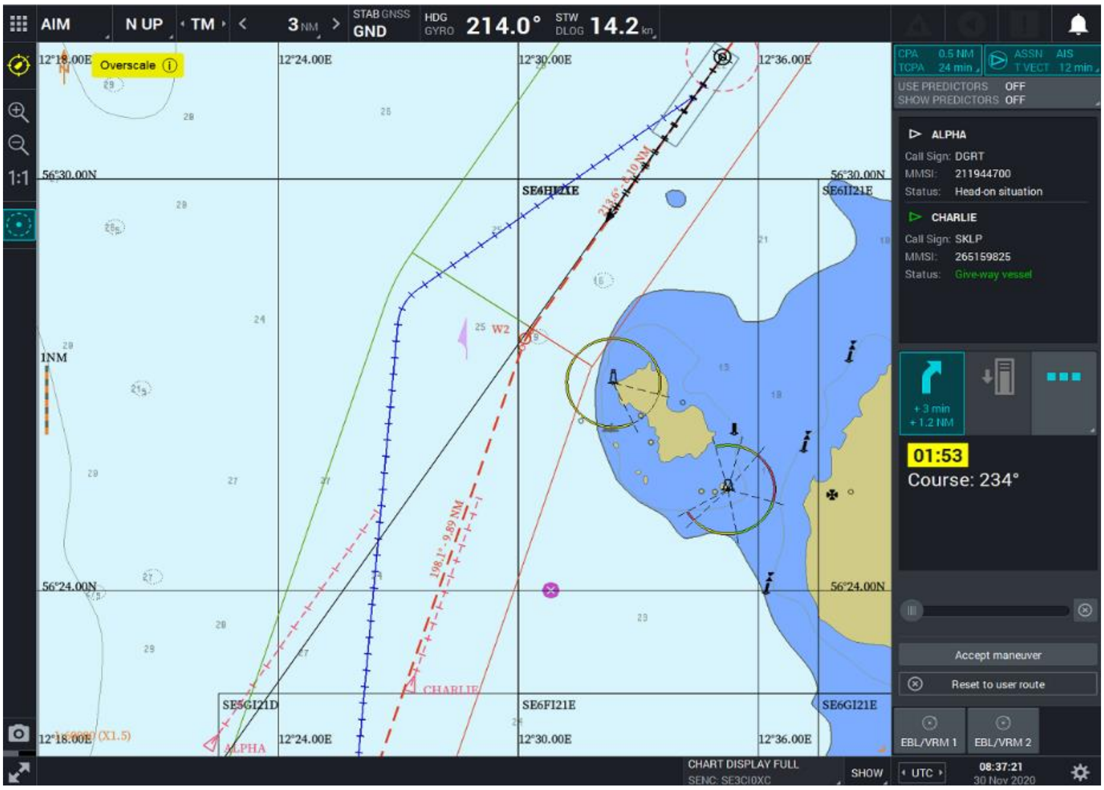


Figure 1 Example of a change of course suggestion (blue line) provided by AIM

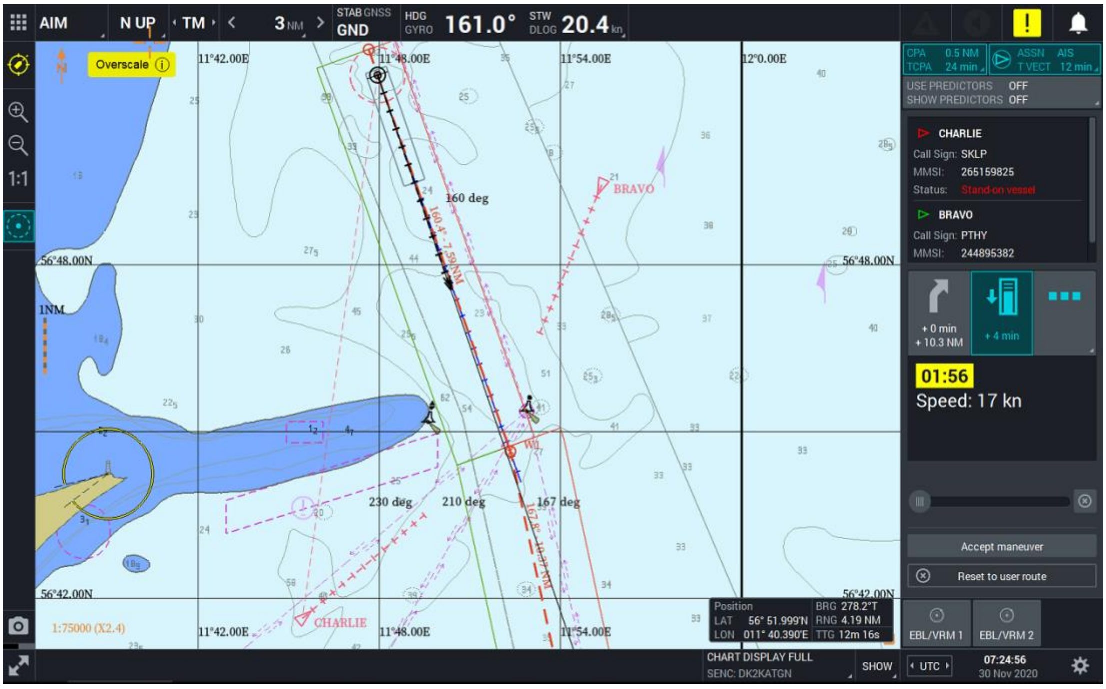


Figure 2 Example of a change of speed suggestion provided by AIM

4.2 Algorithms used in AIM

For reasons of property rights and commercial considerations, Wärtsilä did not disclose the algorithms of the decision support system in any detail but only stated that it is based on COLREGs, anti-grounding and normal behaviour of ships according to statistics based on piloted research (Wärtsilä 2020). According to Wärtsilä AIM presents suggested manoeuvers based on (Wärtsilä 2019):

- The application of the COLREGs based on all identified vessels (AIS, ARPA, other connected sensors), including their course, speed, and navigational status as received by AIS.
- The nautical chart information.
- The ship's route.
- The maneuvering capabilities of the vessel i.e., ship dimensions, max. speed, stopping/acceleration values, ship loading and turn parameters.

4.3 Integration of AIM on the bridge

The AIM software was installed on a separate workstation which was networked to the ECDIS and other sensors and supplied with the following data:

- Compass/gyro, Log, GPS
- AIS
- ARPA

Due to technical reasons ARPA/Radar data could not be integrated in the AIM stations. However, as there was no wind or current in the simulation scenarios, AIS data was considered as being acceptable although not compliant with the COLREGs.

4.4 Operator settings in AIM directly affecting the suggestions provided

In addition to the data provided by the ECDIS and other sensors described in 4.3, the operator can change the CPA/TCPA and the action time settings which directly affects the suggestion provided by AIM for an avoiding manoeuver. Similar to the CPA/TCPA settings in an ARPA, AIM gives a warning for targets within the settings i.e., they are depicted in red, but in contrast to ARPA, AIM also provides a suggestion or several suggestions on how to solve the situation i.e., either change of course/route or change of speed. The CPA value in AIM is considered the ship's safety domain and may be depicted as a circle whereas the TCPA value may be considered as the timing device for when the operator receives a suggestion. If AIM is not able to calculate and present a suggestion within the CPA/TCPA parameters, it will automatically reduce the CPA to a lower value and notify the operator.

The implication of the TCPA setting is that the higher the value, the earlier a suggestion is provided and the higher the CPA value, the more distinct the manoeuver suggestion will likely be i.e., bigger course or speed change. The importance of the CPA/TCPA settings is exemplified in a crossing situation in open waters with equal CPA (depicted as a circle around the ship) but different TCPA settings. With a large TCPA value, give way ship A will receive an early suggestion with a relatively small course change to avoid stand on ship B (Figure 3). In contrast, a small TCPA value will result in a later suggestion involving a bigger course change (Figure 4).

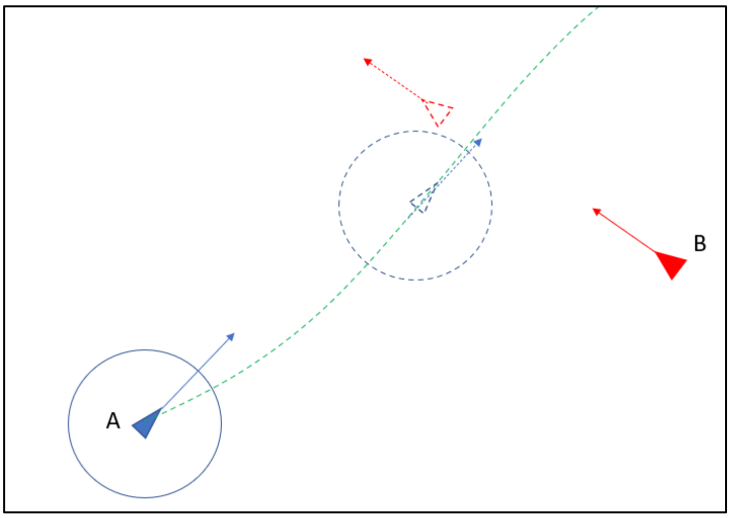


Figure 3 Effect of large TCPA value on AIM suggestion

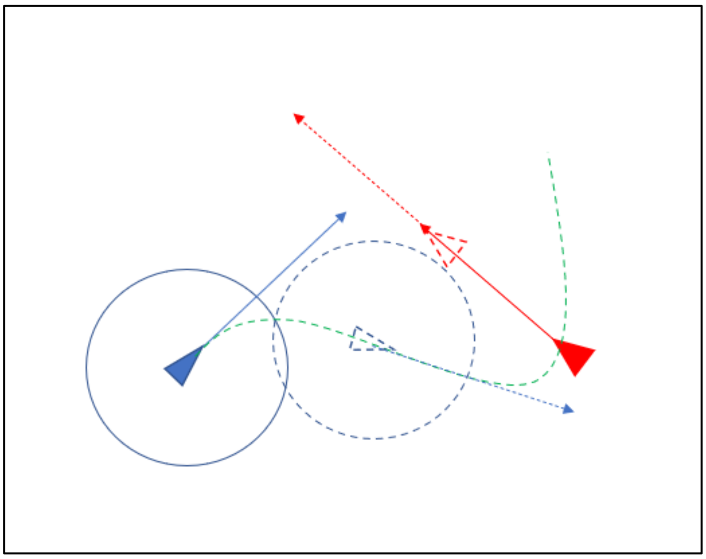


Figure 4 Effect of small TCPA value on AIM suggestion

COLREG Rule 8 states that any action taken to avoid collisions shall, if the circumstances of the case admit, be positive, made in ample time and with due regard to the observance of good seamanship. By being able to adjust the CPA, TCPA and Action Time, AIM does provide some means to actively set values which may reflect the operator's interpretation of at least "positive" and "ample time".

The action time setting is the minimum time after which the operator should perform the avoiding manoeuver. Figure 5 illustrate the parameters, where action time is set to 2 minutes, CPA to 0.5NM and TCPA to 24 minutes (a possible suggestion). As soon as other ships are withing the CPA/TCPA parameters, AIM starts calculating and provides a solution within a couple of minutes. As the Action time is set to minimum 2 minutes, the suggested course change in this case is to be executed after 2 minutes (each dash on the suggested route is 1 minute). However, the Action time setting does not necessarily mean that the suggested manoeuver is to be performed at this minimum time but may be later (see Figure 6 depicting the alternative suggestion 2 in the same traffic scenario). Different settings will provide different solutions i.e., to phrase it differently, the algorithm is not fully automatic but depends on user settings.

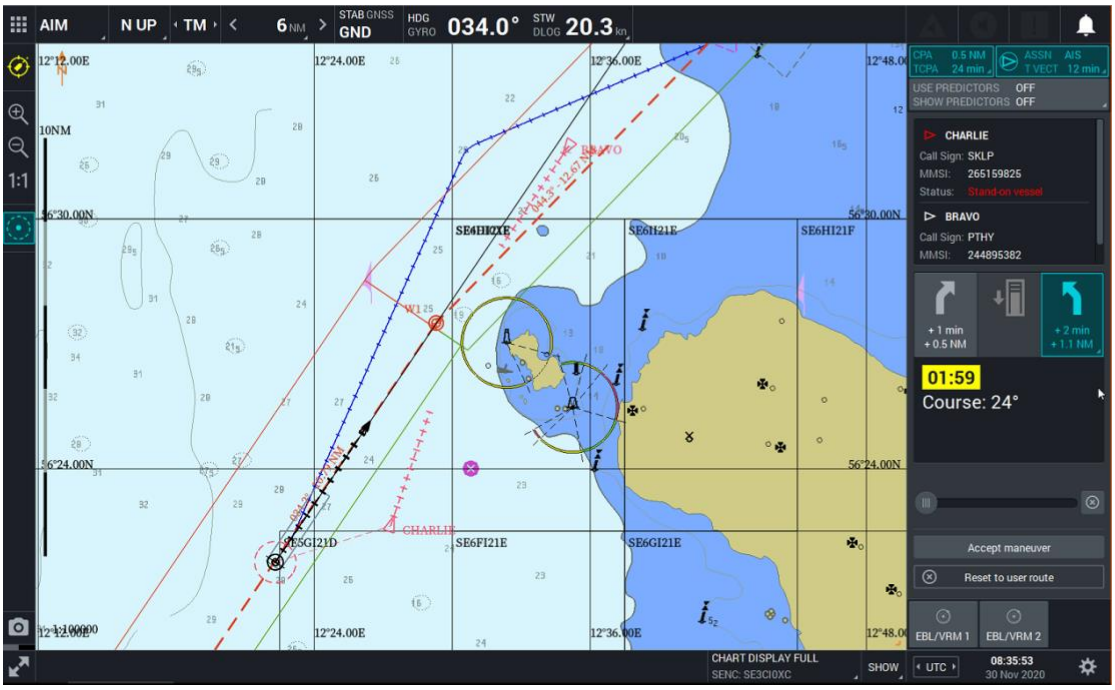


Figure 5 CPA/TCPA settings and Action time with suggestion 1

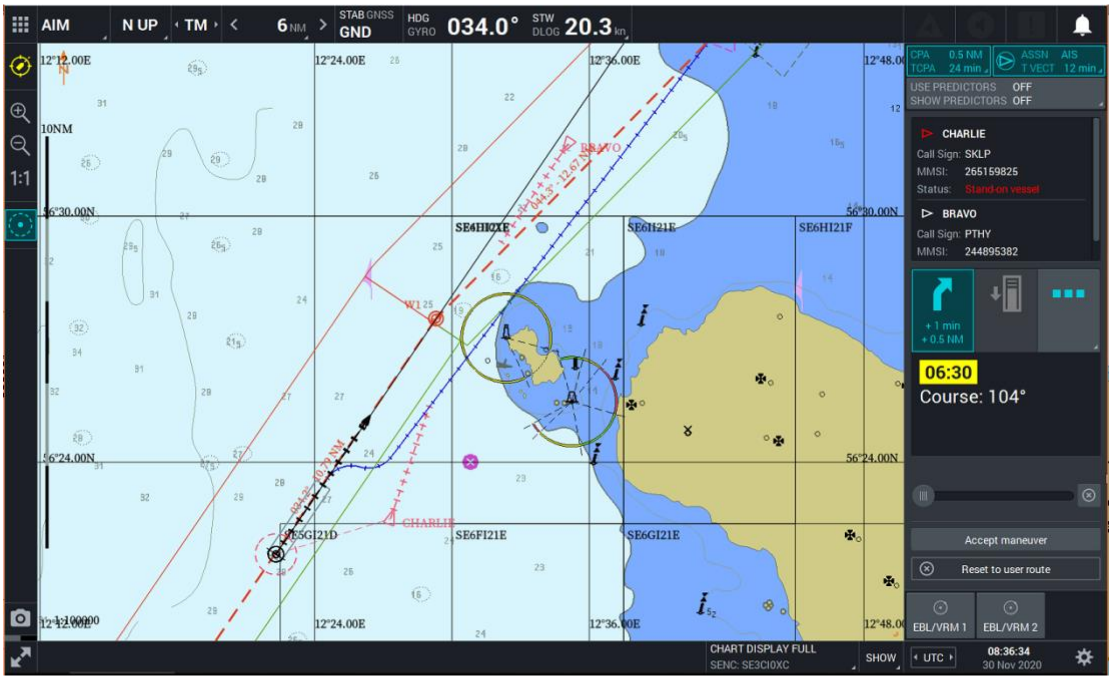


Figure 6 CPA/TCPA setting and Action time with suggestion 2

If the operator does not choose to follow any suggestions, AIM will continuously provide new suggestions which may obviously be rather different from the initial one (see Figures 7 and 8).

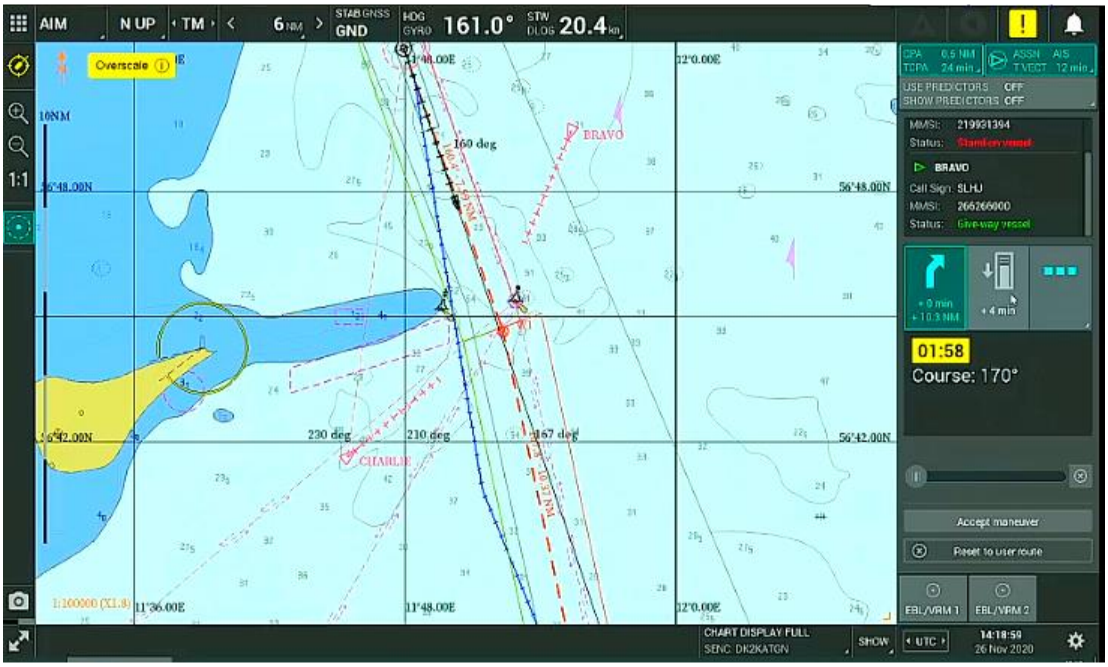


Figure 7 AIM suggestion at start of exercise

Figure 8 AIM suggestion at later stage in exercise

4.5 AIM primary and alternative suggestion(s)

The suggestion presented automatically (i.e., without any operator involvement) as the highlighted blue tile on the right side of the screen is the primary or default suggestion and would be the manoeuvre performed if the ship were autonomous. The dashed red line indicates for which TG the manoeuvre is suggested for, however, the suggestion will “avoid” both TGs. Figure 9 provides an example of the default manoeuvre which consists of a speed reduction to 17 knots and the dashed line points to the main target.

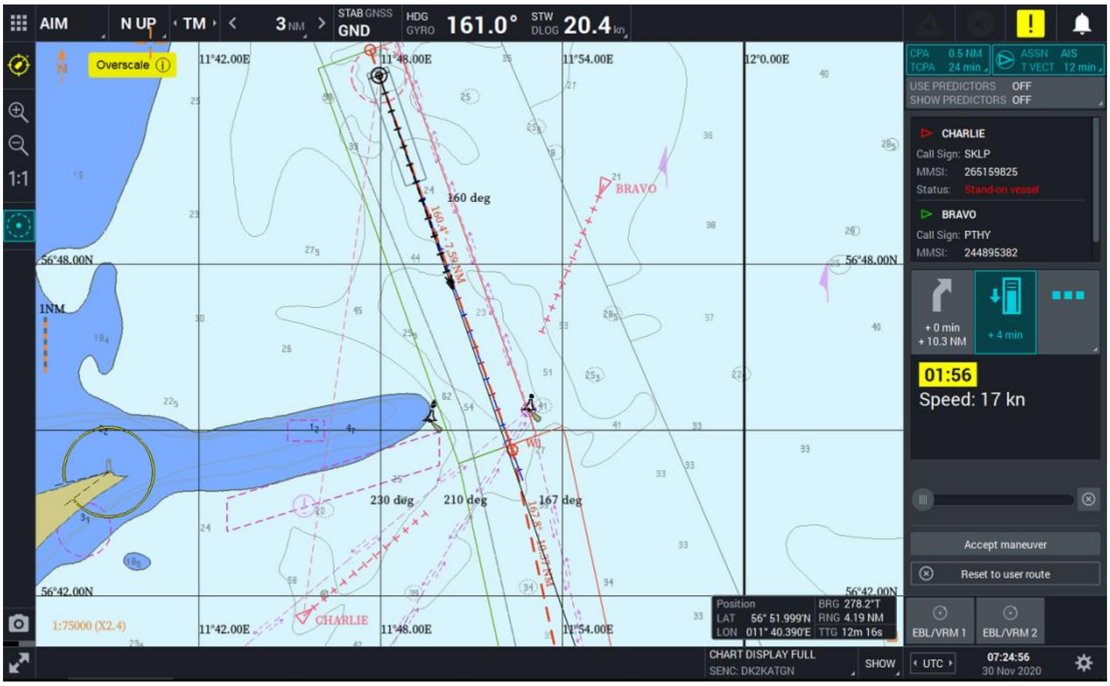


Figure 9 AIM Primary/default suggestion

In this case the alternative solution(s) can be presented to the operator by selecting the respective tile to, in this particular case, the left or right of the primary/default alternative. The tile showing an actual manoeuvre, (i.e., the one to the left) is the second alternative and the tile to the right is the third alternative which may show only three dots (Figures 10 and 11). When clicking on this tile several more options may appear such as turning to port (left), full turn to starboard (right).

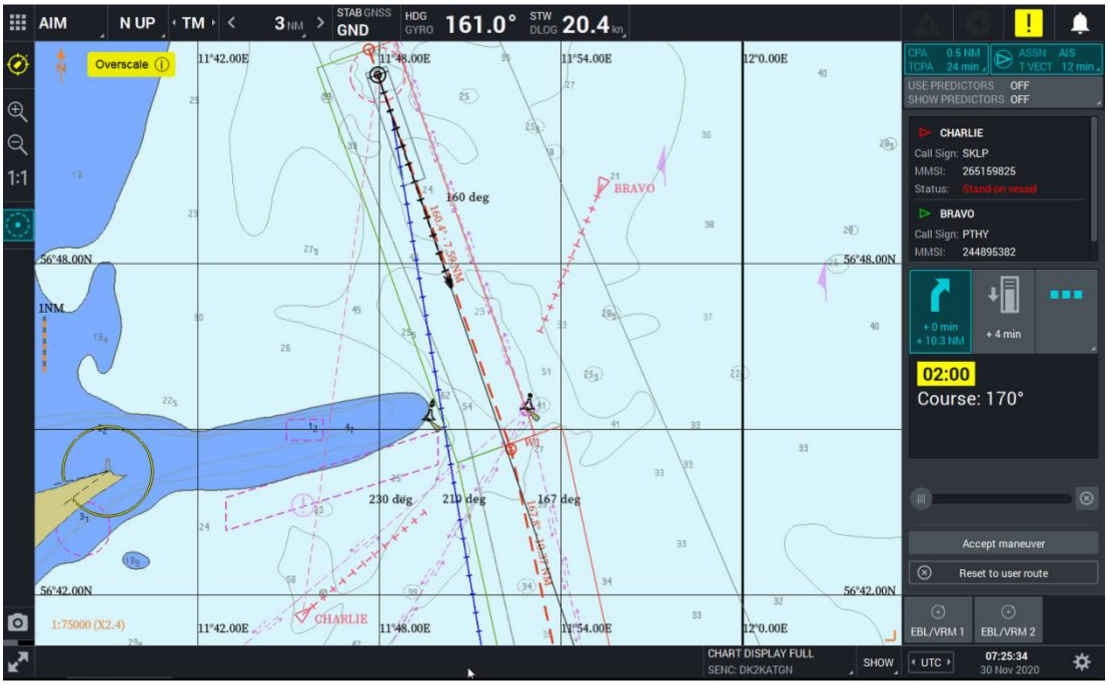


Figure 10 AIM alternative suggestion 1

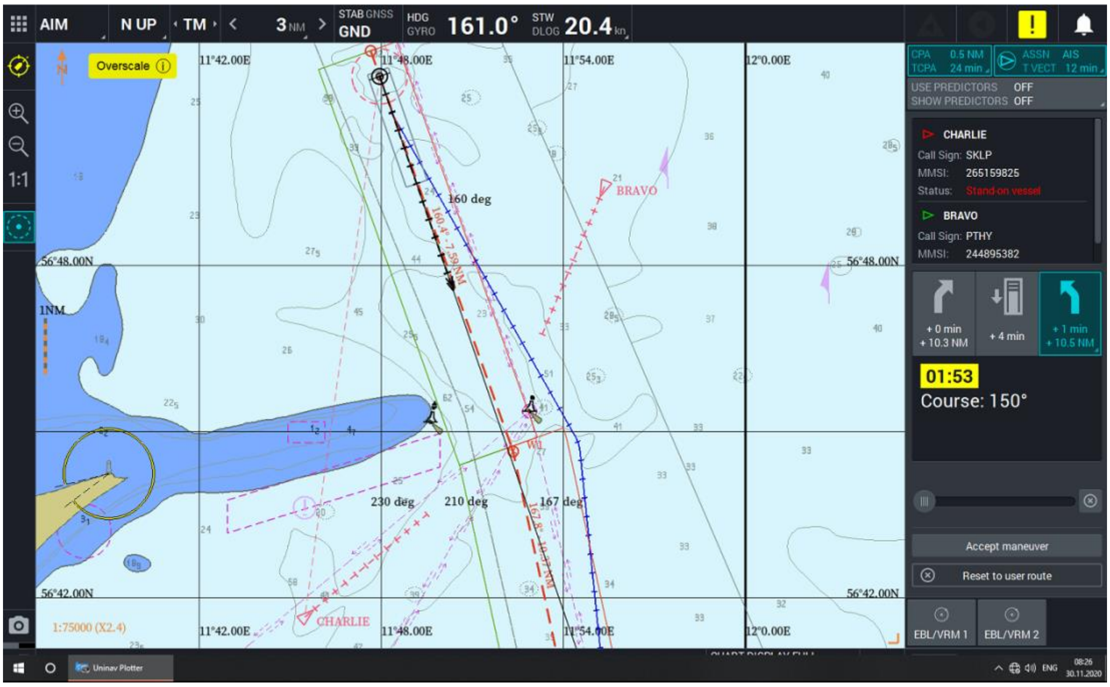


Figure 11 AIM alternative suggestion 2

4.6 "Play ahead" function

AIM allows the operator to virtually execute any suggested manoeuvre by playing ahead to reveal how the OS would end up if following the suggestion provided the TGs keeping their present course and speed (Figure 12).

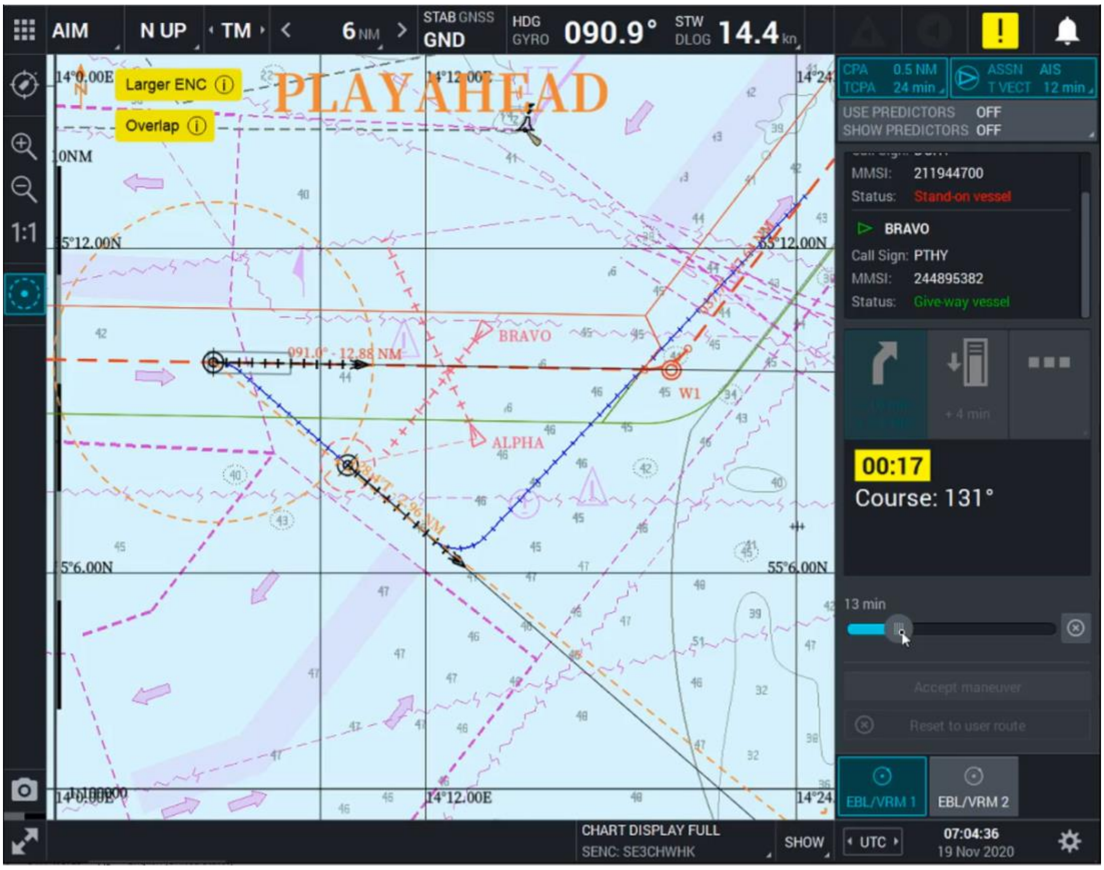


Figure 12 AIM Play ahead

4.7 Executing the suggestion

The suggested route segment provided by AIM can be activated by accepting the manoeuvre and becomes a new monitored route (see Figures 13 and 14). The segment is checked by AIM for navigational hazards with a Cross Track Distance XTD of 0.1 NM on each side of the route meaning that the route segment has been checked for nautical dangers within a corridor of 0.2 NM. If the ship is using track mode autopilot, the course change will be executed automatically once the suggestion is accepted. Suggestions regarding speed need to be executed manually by the operator unless AIM is possibly connected to an automatic speed pilot.

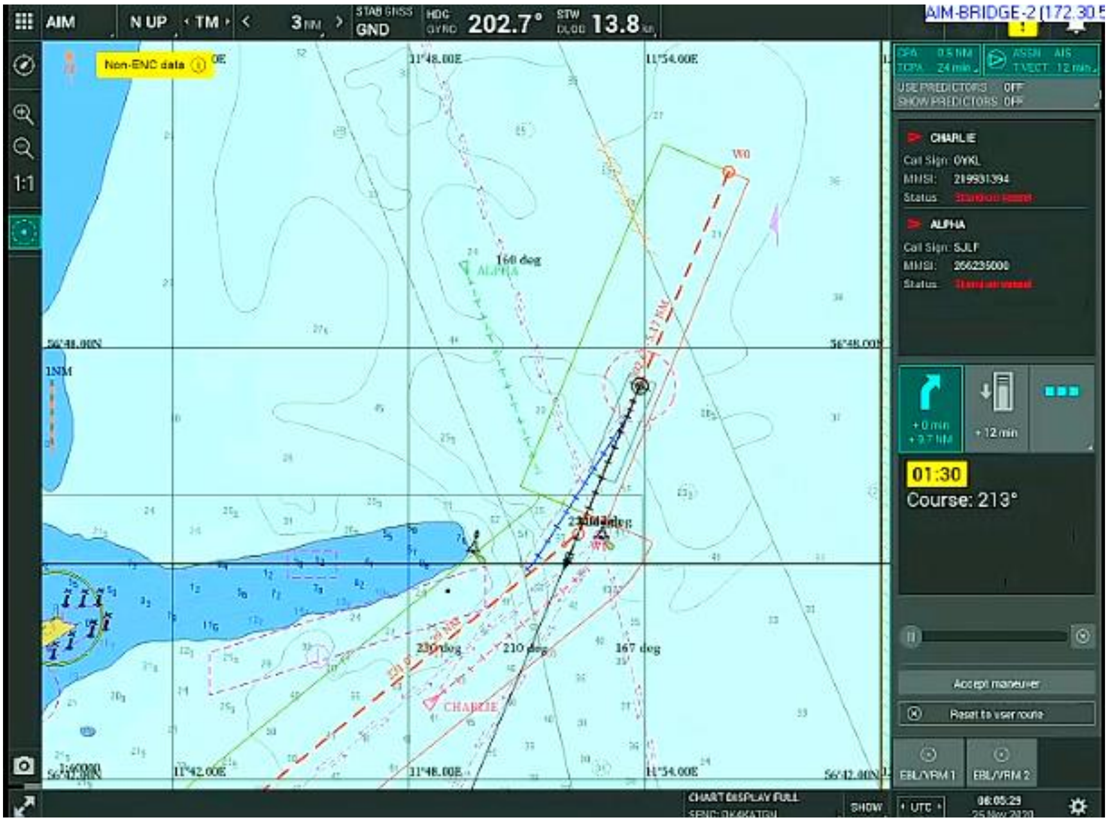


Figure 13 Suggested route

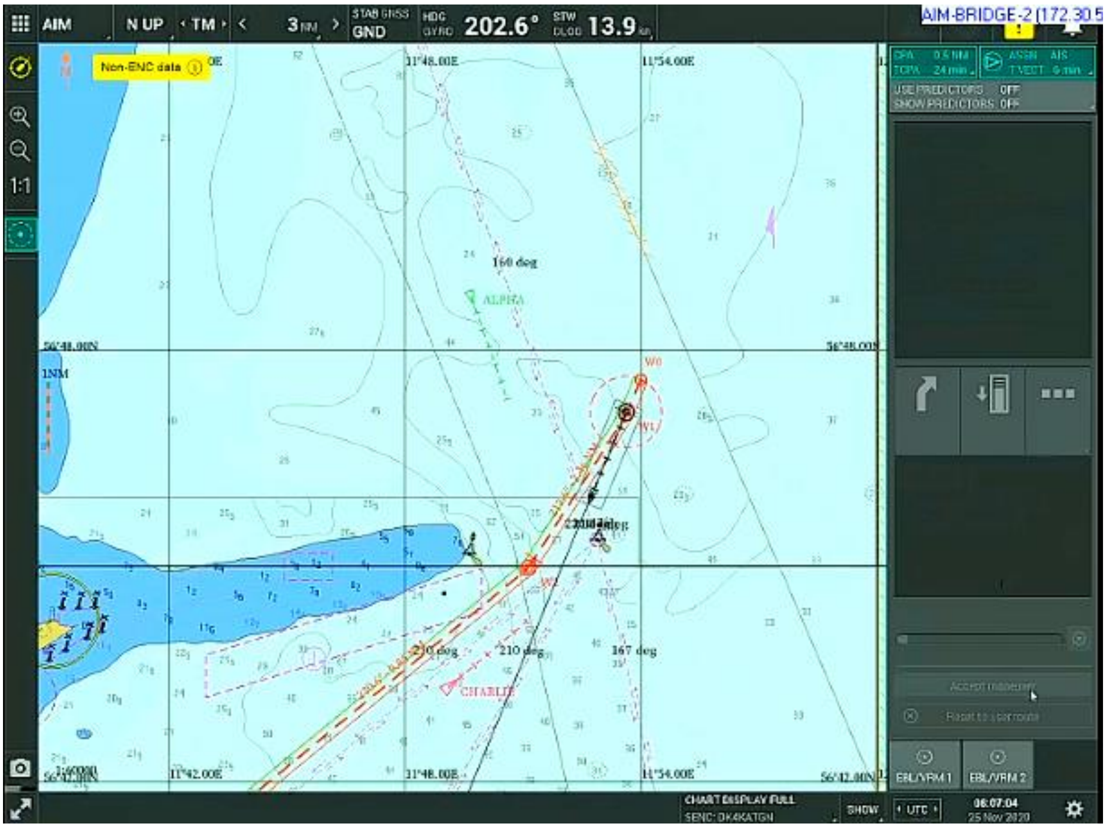


Figure 14 Suggested route accepted

5 Implementation of AIM in Chalmers Full Mission Bridge Simulator

5.1 NTPro simulator software

The AIM software was installed on six separate workstations networked to the simulator server built on Wärtsilä's NTPro 5000 5.40 software. The simulator configuration was set to transmit the ship's data (GPS, log, AIS, etc.) of respective OS/bridge to the AIM station. Also, the route planned on the instructor station for respective ship and bridge was exchanged with the AIM station. No radar data were transmitted to the AIM stations.

5.2 Bridges

Chalmers Full Mission Bridge Simulator (FMBS) was used as a test bed for this research. Chalmers FMBS bridges have a cockpit design with modern integrated navigational (ARPA, ECDIS) and communication equipment (VHF) which were used in this study. Three stand-alone workstations running AIM were installed and configured on the three Chalmers physical bridges. Designated monitors for AIM were placed on each bridge close to the seat from which the test person was navigating the ship (Figure 15).

Figure 15 Bridge 2 set up

5.3 AIM configuration and settings in simulator trials

5.3.1 Ship data settings

Each AIM station on the bridges was configured with the relevant data for the uploaded ship model in the simulation scenario. The data consisted of e.g., manoeuver related data such as dimensions of the ship, deadweight, minimum turning radius, etc. and AIS static data (see as an example the configuration menu of AIM in Figures 16 and 17).

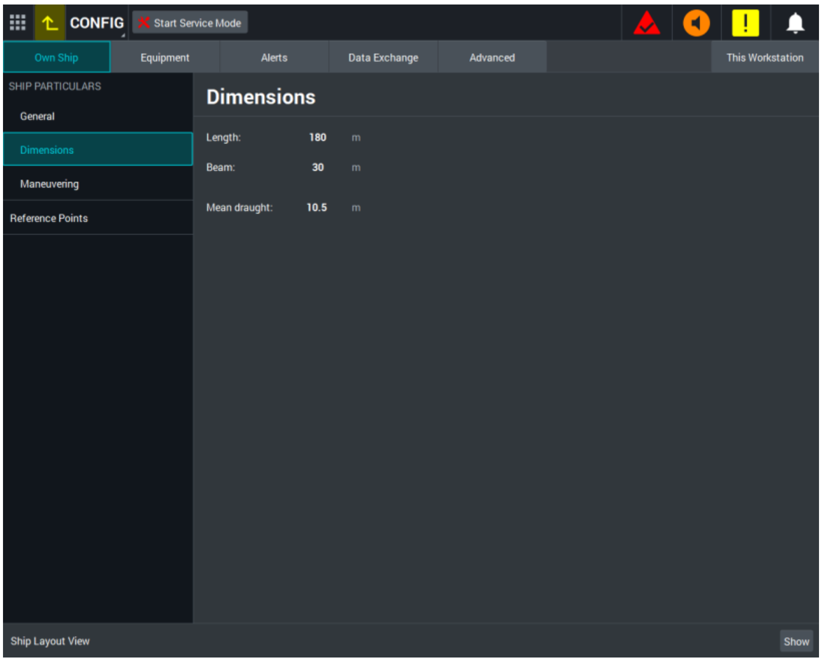


Figure 16 Configuration menu AIM, example dimensions

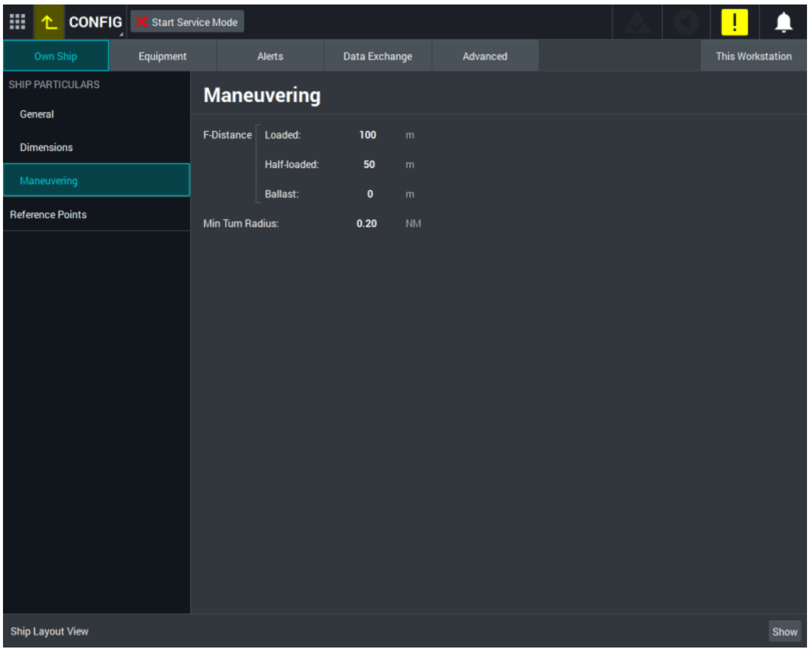


Figure 17 Configuration menu AIM, example manoeuvring

5.3.2 Routes

Each route and corresponding safety settings for each ship and scenario were preplanned by Chalmers researchers using Wärtsilä's NaviPlanner and imported to the simulator software to enable the route exchange between the simulator software and the AIM stations. The route shown on the instructor screen was identical with the route on the AIM stations and the ECDIS on the bridge.

5.3.3 CPA/TCPA settings

As described in section 4.4, the settings of the CPA/TCPA values are influencing the suggestions provided by AIM. To evaluate reasonable settings in the simulation scenarios the following was taken into consideration:

- Exercise geographical area
- Ship types involved in the simulated traffic scenarios
- Samples of real-life values regarding CPA and avoiding manoeuvers based on AIS data
- Review of academic articles relating to CPA and TCPA
- Discussions with subject matter experts
- Large enough TCPA values for AIM to show the suggestions at the start of the exercise

There are numerous papers addressing the actual safety distance/ship domain in real world scenarios. Hörteborn et.al (2019) revisited the definition of the ship domain based on AIS analysis and based on the analysis of 600 000 encounters in 36 locations concluded that the ship domain is formed as an ellipse in overtaking situations, and in crossing situations the ship domain is shaped as a circle where the length of the axes is area dependent (Hörteborn, Ringsberg et al. 2019). According to the paper, the ship domain at e.g., the turn junction off Anholt in open waters is an ellipse with the major axis about 1 NM ahead and astern of the ship and the minor axis roughly 0.4 NM on port and starboard. In confined waters such as the Sound between Helsingborg and Helsingör, the ship domain is rather closer to a circle with an approximate radius of 0.3 NM. Considering the simulation exercise areas and further discussed with subject matter experts, a CPA setting in AIM of 0.5 NM was regarded as appropriate for all exercise areas.

Few studies address real-world close quarters situation and when and at what distance ships take action. AIS data may be used but this data has the disadvantage that it does not give any indication about the monitored route, the weather, visibility and sea state conditions, other ships without AIS in the vicinity, etc. In the case of complex traffic scenarios with more than 2 ships involved or generally dense traffic it may even be unclear as to which target ship the action taken by one ship really applies. Analysis of AIS data at the Anholt junction showed the challenges when using AIS data, where an isolated situation between 2 ships may be easily analyzed (Figure 18), but where multiple ships encounters pose a serious challenge to determine which action on which ship at which distance was taken for which other ship (Figure 19).

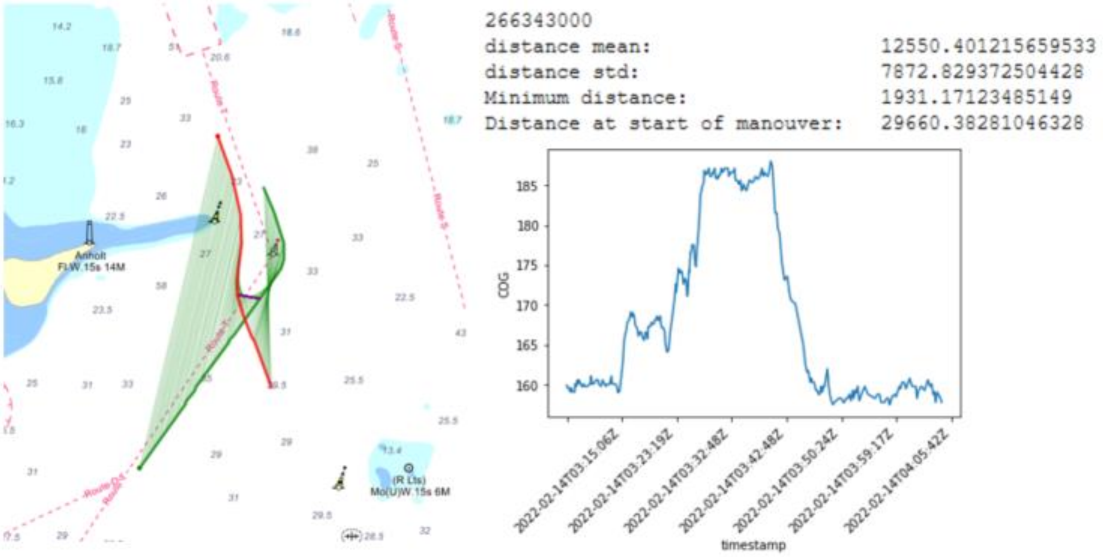


Figure 18 Simple traffic situation between at Anholt junction

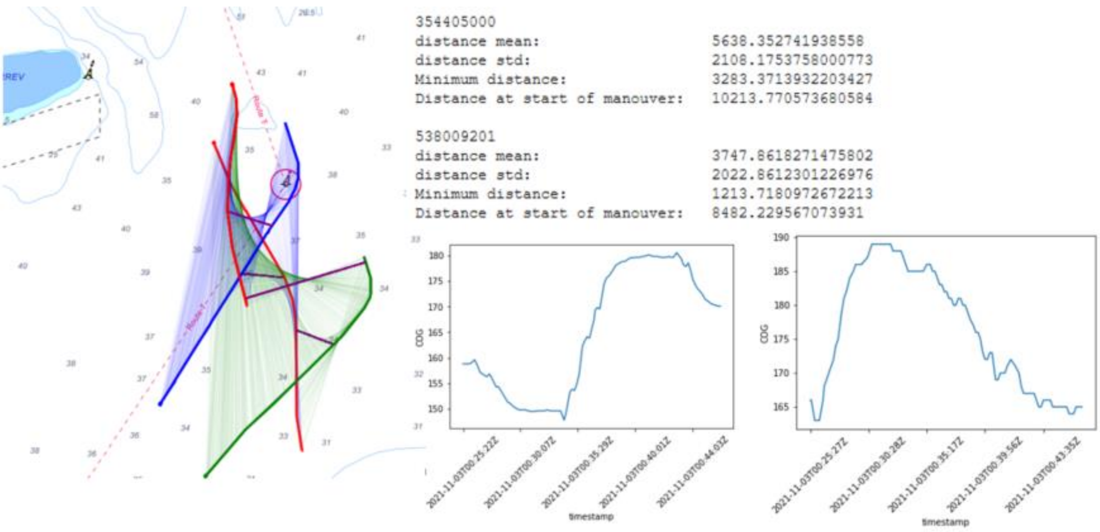


Figure 19 Complex traffic situation at Anholt junction

Eventually, based on the planned scenario length of about 20 to 25 minutes (to be able to run as many simulations as possible in the available amount of time) and that the suggestion(s) provided by AIM should be presented at the start of the exercise, an AIM TCPA value of 24 minutes was chosen.

A paper published after the simulations confirmed that the CPA/TCPA settings in AIM were fairly accurate given the exercise types and areas. Vestre et al (2021) analysis of real collision avoidance manoeuvers based on AIS transmissions from 13 days covering the Norwegian exclusive economic zone identify that TCPA mean value when taking action ranges between about 18 to about 20 minutes in overtaking crossing and head-on situations and result in a mean passing distance of about 1.2 km (0.6 NM) (Vestre, Bakdi et al. 2021). However, the study also showed a large distribution of the data which showed that there is a big variance in both CPA and distance when ships initiate their manoeuvers.

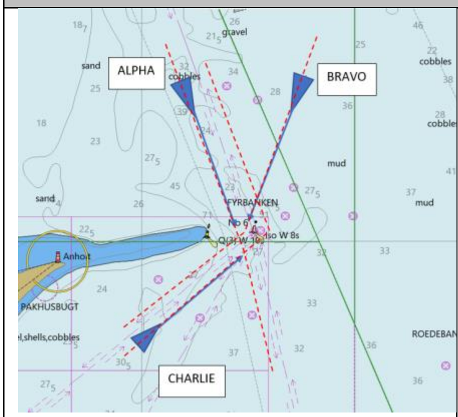
5.3.4 Action time

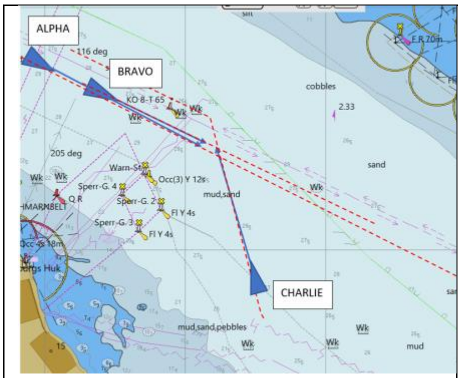
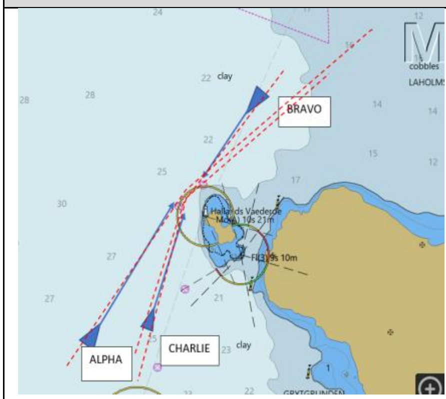
The Action time was discussed and evaluated by subject matter experts to correspond to reasonable setting of 2 minutes.

6 Simulator Scenario development

Subject matter experts (SME) were involved in creating, implementing, and testing various traffic scenarios in the Wärtsilä NTPro 5000 Full Mission Bridge Simulator (FMBS). The goal was to develop scenarios that were realistic and somewhat challenging for both the participants and the AIM software. The scenarios had to meet the following criteria: include meeting, overtaking, and crossing situations, occur in neither open sea nor restricted waters, good visibility, calm weather conditions, manageable for one single officer on the bridge, and have a duration of approximately 20 to 25 minutes, allowing the test person enough time to assess and act upon a situation. A total number of 10 scenarios were developed and pilot tested, of which three satisfied the pre-conditions. All scenarios involved three vessels, Alpha, Bravo and Charlie and were set in three different geographical areas: Anholt, Fehmarn, and Halland (Table 1).

Table 1 Simulator scenarios

Scenario	**Scenario Description**	**Distance between ships**			
Anholt			**Alpha**	**Bravo**	**Charlie**
	According to the COLREGs BRAVO is crossing in relation to ALPHA and therefore give-way vessel. ALPHA is give-way vessel regarding CHARLIE but is limited in manoeuvring space to starboard. CHARLIE is stand-on vessel to both ALPHA and BRAVO and shall keep course and speed.	Alpha	X	4,6NM	10NM
		Bravo	4.6NM	X	9.7NM
		Charlie	10NM	9.7NM	X
Fehmarn			**Alpha**	**Bravo**	**Charlie**

	ALPHA is overtaking BRAVO, and both are give-way ship for CHARLIE. Manoeuvring space to the starboard side of both ALPHA and BRAVO is limited due to a prohibited military area marked by special marks.	Alpha	X	2.1NM	10.9NM
		Bravo	2.1NM	X	8.9NM
		Charlie	10.9NM	8.9NM	X
Halland			**Alpha**	**Bravo**	**Charlie**
	ALPHA is overtaking CHARLIE and in a head-on situation with BRAVO. All ships have limited manoeuvring space to their East. CHARLIE is stand on vessel in regard to ALPHA and give way vessel in regard to BRAVO	Alpha	X	12.3NM	2.9NM
		Bravo	12.3NM	X	10.2NM
		Charlie	2.9NM	10.2NM	X

7 Methodology

This study adopted a pragmatic mixed methods approach to address the research questions. Although there is a heavier reliance on qualitative data, the inclusion of quantitative data afforded an opportunity to understand the problem more clearly, resulting in a more complete analysis (Creswell and Clark 2017).

The development of this work follows a convergent design which is when the researcher obtains different but complementary data on the same research topic, to address the strengths and weaknesses of both qualitative and quantitative methods (Creswell and Clark 2017). The convergent mixed methods design procedure was followed to develop and complete this study. First, research questions were identified, followed by the identification of the appropriate approach for each method which included operationalizing the variables, defining the data collection methods, and finally a reflection on the quality of the data collection methods. The convergent design was selected as the most appropriate method for this work given the limitations associated with the data collection. Due to the COVID-19 pandemic, there were time constraints, financial limitations, participant recruitment difficulties, and university implemented restrictions.

8 Participants

Participants were fully informed of the procedures and risks of the experiment and signed written informed consent prior to the start of the data collection. The experiment complied with the requirements of Article 28 of the EU General Data Protection Regulation (2016/679) regarding protection for physical persons in the processing of personal data. Each participant was assigned a unique identification number (ID) prior to arrival, which was used for the questionnaires throughout the study to maintain confidentiality. Purposive sampling is a non-random technique used when the participants need to have certain qualities, skills, knowledge or experience (Etikan, Musa et al. 2016). Purposive sampling was used to recruit professional mariners (active or recently active masters, mates, officers, maritime pilots, or fourth year Master Mariner students) as test participants. Participants were recruited through various social media platforms, Chalmers professional maritime network and word of mouth. In total nineteen participants were recruited. There was no financial compensation for participants in the study.

A demographic survey was completed when the participants arrived. All nineteen participants were of Swedish nationality and identified as male. In addition to basic demographic information, they were also asked about their attitudes towards automation through receiving decision support through technical means in a navigational situation. See Table 2 for a summary of the demographic information.

Table 2: Summary of demographic information

Category	Sub-category	Frequency
Age	18-24	6
	25-34	4
	35-44	2
	45-54	4
	55-64	3
Current Position	Fourth year master Mariner student	7
	Not currently employed	1
	Employed as OOW/Senior Officer	3
	Other * (2 pilots, 1 director, and 1 researcher)	4
	Chalmers instructor or teacher	4
Attitude towards receiving decision support for navigation	Very negative	0
	Negative	0
	Neutral	4
	Positive	13
	Very positive	2

9 Experimental design

The experiment lasted for approximately four hours per session, including a technical familiarization period and the group interviews at the end of the day. The study began with familiarization, consent, and demographics questionnaire. Next, the scenario was uploaded on the bridges and set to pause mode. Participants were assigned to a bridge (Alpha, Bravo, or Charlie) and each participant was given approximately 10 minutes to complete the pre-scenario questionnaire which was an assessment of the situation including plan of action and expectations of the other ships in the scenario. Then, the instructor switched on AIM (if applicable), allowing the participants time to study the suggestions. The exercise began and lasted for approximately 25 minutes at which point the participants completed the post-scenario questionnaire. The same procedure was repeated for each exercise. After three scenarios were completed, the participants and research team returned to the classroom for a group debriefing and interview.

This study followed a within-subject design. All participants were exposed to both conditions: baseline condition (traditional navigation) and a decision-support condition (navigation with AIM). However, they were never exposed to the same scenario in both conditions. The order in which the scenarios were tested, along with the condition of baseline or AIM were randomized to reduce any potential order effect (Creswell and Clark 2017). There were an uneven number of trials between baseline and AIM due to last minute COVID-19 cancellations. Table 3 provides a summary of the experimental conditions.

Table 3: Breakdown of experimental conditions

Scenario	**Baseline**	**AIM**
Anholt	4	3
Fehmarn	3	4
Halland	3	4
Participants (x3 per scenario)		
Total Trials	30	33

10 Data Collection

The data were collected during the period of November-December 2020. Qualitative data were collected through a pre-scenario questionnaire, video and audio recordings, observations, and the group interviews. The questionnaire was first completed during the pilot study and adapted as necessary after it was evaluated by two subject matter experts to ensure face and content validity. The pre-scenario questionnaire was intended to represent a "watch takeover" which occurs at the end of a watch, the Officer of the watch (OOW) passes the navigation of the vessel to the relieving officer. This is an important aspect of navigation and requires careful review of positional and traffic information to ensure a safe passage. The participants were given ten minutes to evaluate the traffic situation based on the available information from their visual observations, ECDIS, ARPA, AIS and AIM (if available) and indicated the:

- primary TG for which action will be taken.
- applicable action planned to be taken.

- approximate distance and or time when the action is planned to be executed.
- predicted behaviour of other ships in the scenario.

The data from the participants were then compared with both their own result (i.e., did they follow their own plan) and with AIM's primary (default) and in certain cases alternative suggestion(s) on respective bridges. Note that only the initial suggestions, at the start of the exercise were analyzed. No subsequent suggestions made by AIM during the exercise itself were compared with the test persons evaluations.

Interviews were conducted after all three scenarios were completed and included all three participants and available researchers. The interviews were semi-structured and lasted for approximately one hour. The interviews were video and audio recorded, and a complete transcription was completed post interview and compiled with the researcher's personal notes made during the interviews. The interview was structured around questions related to the simulation, overall experience, interaction with AIM, and perception of COLREG adherence and seamanship.

Quantitative data were collected from the simulator generated log files and AIM route log files complemented with AIM and radar screen recordings. NTPro software provides the user with automatically generated log files for each exercise including positions, speeds, position of targets, and rudder angles.

VHF ship to ship communications were recorded using the bridges audio/video recording software but almost no communications occurred during the simulations, therefore this measure was disregarded.

11 Data Analyses

The data analyses were completed in two steps. The first step involved a review of the pre- and post-scenario questionnaire responses, compared to the AIM suggestions and the ship track logs. The following analyses were completed:

- Comparing the identification of the primary TG by the navigator and the primary TG according to AIM.
- Comparing the planned actions of a navigator with AIM suggestion(s).
- Comparing the expectations of a navigator on what other ships will do in the traffic situation (predicted behaviour) with the AIM suggestion(s) on the other ships.
- Comparing and contrasting the participants' planned actions and predicted behaviour with the actual execution.

- Identifying where AIM is possibly influencing the actions/behaviour of a navigator, e.g., situations where the participant has planned a manoeuver and AIM suggested a different one.

The group interviews were analyzed by two independent researchers to achieve intercoder reliability to improve the transparency, and trustworthiness of the analysis (Patton 2002, SAGE 2008). The researchers each followed the same general process to complete a thematic analysis (Braun and Clarke 2006). Once the individual analysis was completed, the researchers compared and discussed the results, generating a common list of the themes that emerged from the data and then any discrepancies were addressed. The researchers continued working with the data and finalized the analysis when saturation was achieved and the themes which emerged from the data were clear.

12 Limitations

- This study was conducted during the COVID-19 pandemic leading to a limited number of participants.
- The participant sample is homogeneous consisting of Swedish, males, with similar experience and training in navigational situations. This may limit the generalizability of the results to other geographical areas and seafaring populations.
- Only three traffic scenarios were tested in this study.
- The repeatability of the results is limited as researchers require access to AIM, which is not yet commercially available. However, the scenarios described, and methods used can and should be implemented for testing similar decision support systems.

13 Results

13.1 Comparison of identification of TG for which the action is taken (primary TG)

Although AIM suggests a manoeuver which will avoid a close quarters situation with any TG within the set parameters, the software also identifies the primary TG for which the manoeuver applies. However, note that the proposed suggestion by AIM is "avoiding" both TGs. Comparing the AIM data with the participant data regarding the primary TG gives potentially an indication of the possible misinterpretations of the COLREGs regarding stand-on and give-way vessels in the scenario. The comparison was classed as either "same as AIM" or "different from AIM".

Table 4: Comparison of Primary Targets

	ALPHA		**BRAVO**		**CHARLIE**		**ALL SHIPS**	
Same as AIM	13	61.9%	17	85.0%	14	93.3%	44	78.6%
Different from AIM	8	38.1%	3	15.0%	1	6.7%	12	21.4%
Total frequency	21		20		15		56	

Table 4 gives an indication if the participant and AIM are planning the manoeuver with the same TG "in mind" and if the COLREGs are applied correctly, e.g., if a TG is considered give way or not. The lower number for ALPHA is probably due to the possibly ambiguous COLREG situation Anholt (BRAVO crossing or being overtaken by ALPHA, see figure 20) and Halland where AIM on ALPHA was taking CHARLIE being overtaken as primary target whereas many participants focused on BRAVO as head on situation (figure 21).

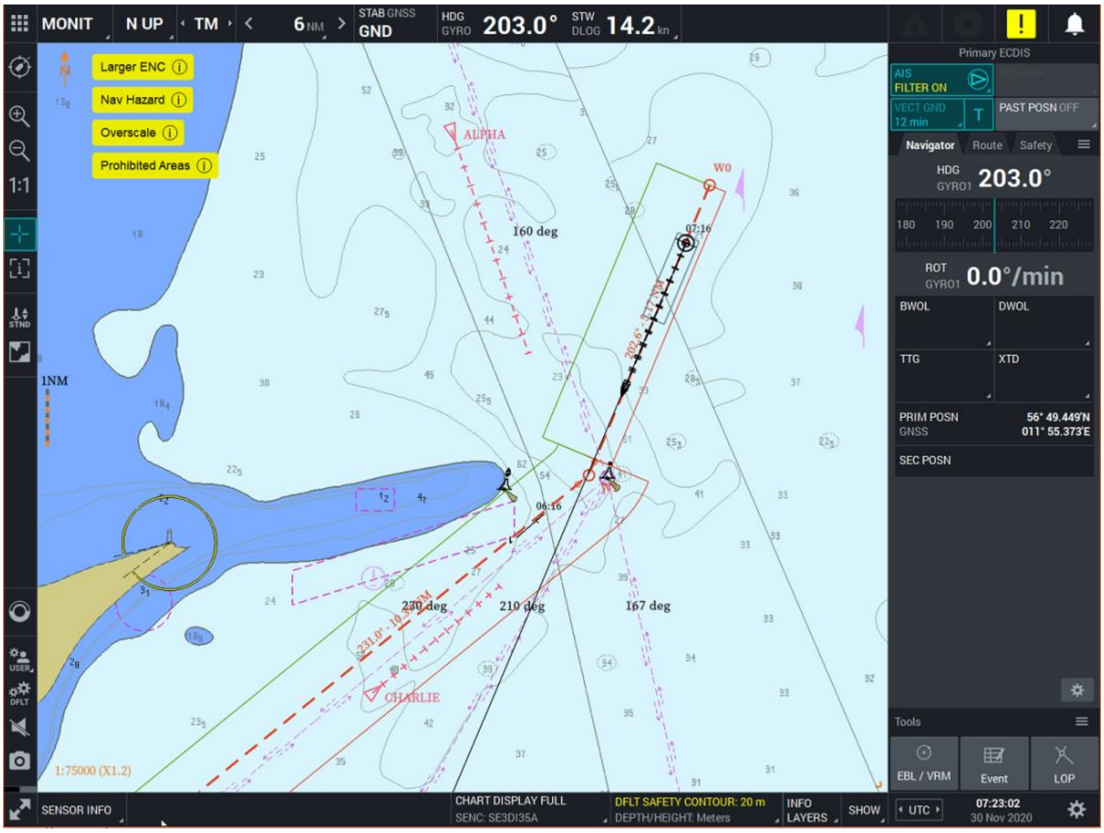


Figure 20 Situation between BRAVO and ALPHA in Anholt scenario

Figure 21 Situation of ALPHA in Halland scenario

It may be stated that AIM is calculating strictly according to the COLREGs where, for example, a clear distinction is made that the ship is considered overtaking or crossing. However, participants may not make the same calculation and therefore may perceive such a situation differently especially if there is a borderline between crossing/overtaking where other factors may influence a participant's judgement. It was assumed, for example, by several participants on ALPHA in the Anholt scenario that ALPHA is overtaking BRAVO and therefore should keep out of the way. Setting the priority target obviously has an influence on the planned manoeuver, i.e., a planned manoeuver applicable for one target may be different than a planned manoeuver for another target. However, in general, the participant's view has a high level of agreement with AIM.

13.2 Association with planned manoeuver of the OS and AIM suggestion(s)

The participant's planned manoeuver to solve the traffic situation was compared with AIM's default/primary suggestion and alternative suggestion(s) to evaluate the similarities and differences of the suggestion(s). The associations between the plan were classified and tabulated into:

- Strong: virtually the same as AIM but may differ slightly (time of execution and amount of course/speed change).
- Moderate: same type of manoeuver as AIM but may have been done later and therefore with bigger speed/course change.
- None: manoeuver not suggested in AIM

The AIM CPA/TCPA value has a significant effect on the algorithm-suggested manoeuver and as the participants' consideration of what is regarded as a safe TCPA may not match with the AIM setting (i.e., participants may act later but with a bigger manoeuver), it was agreed to group "strong association" and "moderate association" together (Table 5).

Table 5: Comparison of planned manoeuvre of the OS and AIM suggestion(s)

	ALL SHIPS					
	Primary		Alternative 1		Alternative 2	
Strong association	17	72,7%	8	30,9%	4	11,6%
Moderate association	23		9		1	
No/Little association	15	27,3%	38	69,1%	38	88,4%
Total frequency	55		55		43	

The summary for all ships and scenarios indicates that AIM's default suggestion is strongly or moderately associated with the participants own plan to quite a large extent. However, on an individual ship basis there are rather big differences which need to be explained.

In situations where one ship is stand-on vessel in relation to one target and give-way to another one, AIM disregards the stand-on obligation and focus on the give-way situation. However, the suggested manoeuver will not lead to a collision with the give way vessel as long as this ship keeps her course and speed. In other words, AIM is always calculating the other ship's position based on a straight vector. Participants seem to plan in a similar way but foresee to make the manoeuver (often the same type as AIM) at a later stage and thereby allow time to ascertain if the situation develops as expected. As the results reflect both strong and moderate associations, the percentage of matching what AIM suggests is quite high. The challenge with AIM is the decision regarding the TCPA value which can be regarded as both individual and situation dependent. As the suggested manoeuver was set with a TCPA of 24 minutes it resulted, in many cases, in a moderate association with the AIM suggestion as the participant's own plan mainly differed in to perform the avoiding manoeuver. At the same time, having a high TCPA value in AIM also allows a timely action to a possibly unexpected solution, e.g., going to port in a head-on situation although at 12 NM distance (Figure 22).

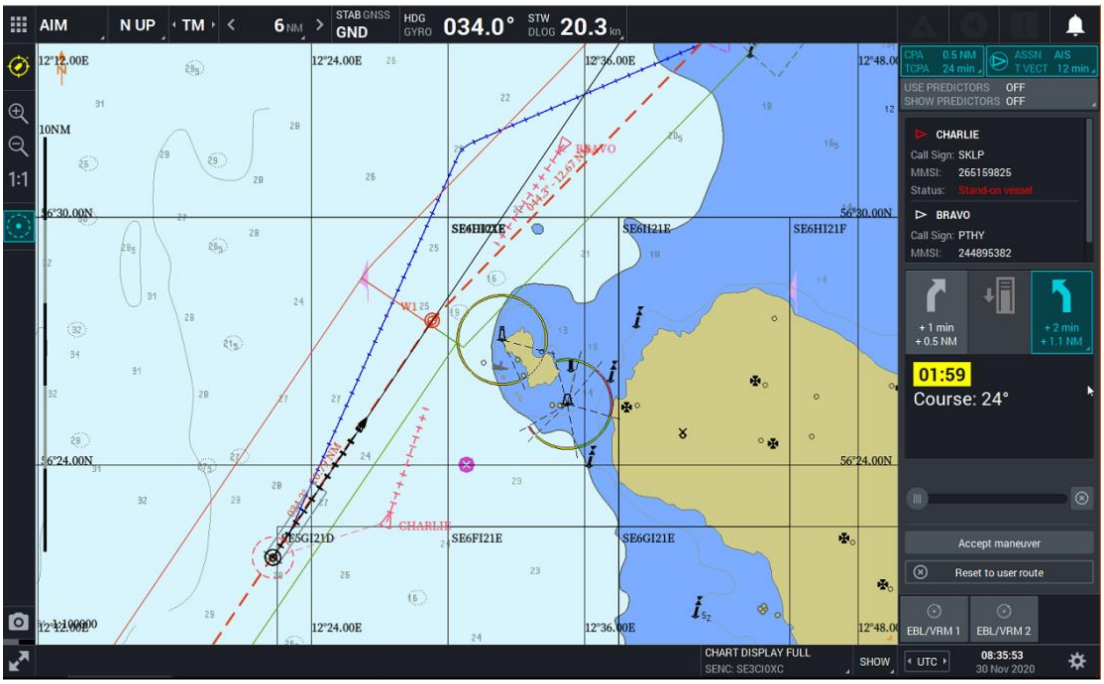


Figure 22 AIM primary suggestion on ALPHA in Halland scenario

This solution may be called "good seamanship" by keeping away from land and away from traffic but was not considered by any participants but one. One would also need to consider the implication of doing such a manoeuver if the other head on ship strictly follows the COLREGs and turns to starboard. Generally, the TCPA setting of AIM is clearly a major factor when trying to match the AIM suggestions with the participants' plans. The more the AIM TCPA setting matches with the participant's mindset or collision avoidance strategy, i.e., if a participant tends to take early action (possibly less distinct) or later (and more distinct), the higher the association between the participants' plans and AIM suggestions become. However, these AIM TCPA settings cannot be set at a fixed number and need to be adapted to among others geographical area and traffic density.

For a purely stand on vessel, e.g., CHARLIE in the Fehmarn scenario, participants seem to have struggled with AIM presenting a suggestion at all as they considered that they must only keep course and speed. However, all AIM suggestions proposed course and speed for at least 15 minutes and the course/speed change after that time needs to be considered as a suggested manoeuver provided that the give way vessels in the scenario do not take any action (as an example, see Figure 23).

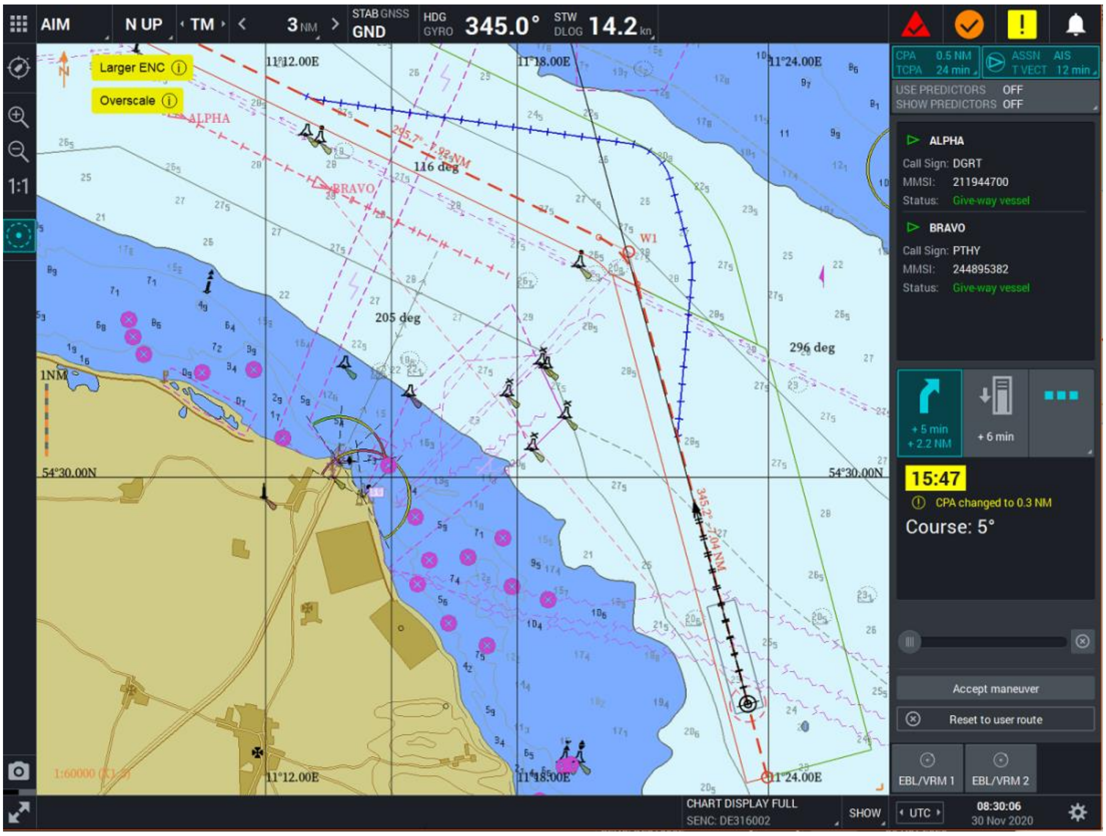


Figure 23 AIM suggestion on CHARLIE in Fehmarn scenario

13.3 Association of OS predicted behaviour of TGs and AIM suggestion(s) on TG bridge

Before starting the exercise, each participant stated how the other ships (TG) are expected to manoeuver to solve the traffic situation. The participant's solution was then compared with the AIM suggestions on the TGs. The purpose was to get an idea how well the AIM suggestions agreed with the expectations of a human operator. Correct prediction means only the same type of manoeuver, e.g., course change to port or starboard, speed reduction or keeping course and speed as AIM suggested on the other ship's bridge but not the magnitude of course changes or speed reductions (Table 6). When comparing these data with "OS plan before starting" only the primary suggestion had a strong/moderate prediction. However, when evaluating how the TGs may manoeuver, the other suggestions are also taken into consideration when possible.

Table 6: OS predicted behaviour of TGs and AIM suggestion(s) on TG bridge

	ALL SHIPS					
	Primary		Alternative 1		Alternative 2	
Correct prediction	56	50.9%	41	37.3%	28	30.1%
Incorrect prediction	54	49.1%	69	62.7%	65	69.9%
Sum	110		110		93	

The main reasons behind the rather low predictiveness for all ships scores may be related to the following reasons:

1. Different interpretation of COLREGs stand on or give way vessel (in case of ALPHA and BRAVO at Anholt, see Figure 20) where some participants on BRAVO incorrectly assumed that ALPHA is overtaking BRAVO and accordingly is a give-way vessel.

2. Different solution to solve the situation which may have been covered in AIM alternative 1 or 2 (e.g., expected behaviour of ALPHA from BRAVO perspective at Anholt and Fehmarn). This would mean that AIM's alternative suggestions may be as equally valid as the default suggestion, according to the participants. This also introduces an interesting question related to how AIM prioritizes the suggestions.

3. Only considering the relation of own ship with e.g., TG1 but not considering the relation TG1 and TG2 may have a critical influence the decision regarding actions with TG1. Although participants, in general, claimed that the relation between targets is important when deciding on a manoeuver, it seems that they did not necessarily consider these interactions and their implications in more detail.

4. Expecting preemptive measures of a stand on vessel (e.g., CHARLIE in Fehmarn scenario to change course to starboard early) which is never suggested by AIM as it is purely calculated based on current course and speed.

The analysis of the level of agreement of the primary AIM suggestion on the TG is of particular interest as this is the manoeuver an autonomous ship would ultimately execute in a given traffic situation. The analysis of the data may give an indication whether traffic situations that have both human operators and smart/autonomous vessels be safely resolved compared to situations encountered today's traffic schemes. By assuming that the risk of a close quarters situation is reduced if the autonomous ship acts as expected by a manned ship i.e., if the primary AIM suggestion on the TG is equal or similar to what the OS expects, the human operator will not even know if the TG is manned or fully autonomous, a Turing Test conundrum!

13.4 Comparison of planned and actual manoeuver

The actual track and speed of each own ship (OS) and run were compared with the initial plan. A strong agreement means that the OS was following the initial plan. The purpose of these data is to evaluate the influence of unforeseen factors during the exercise and/or the possible influence of AIM on the initial plan (Table 7).

Table 7: Comparison of planned and actual manoeuvre

	ALPHA		**BRAVO**		**CHARLIE**		**ALL SHIPS**	
Strong Agreement	12	85.0%	10	90.5%	11	88.2%	33	87.9%
Moderate Agreement	5		9		4		18	
No Agreement	3	15.0%	2	9.5%	2	11.8%	7	12.1%
Sum	20		21		17		58	

The data indicates that participants mostly adhered to their initial plan with some adaptations in 30% of the cases. Only in 7 cases they made a different manoeuver than planned which could be due to AIM (if AIM was available) or due to the non-expected behaviour of the other manned ships. However, it is interesting to note that in 18 of 58 cases (31%) the planned manoeuver needed some adaption as it was only moderately consistent with the original plan. Also, although participants incorrectly anticipated the behaviour of at least one of the other ships in the scenario in 48.4% of the cases, this did not significantly influence the execution of their plan.

13.5 Comparison of predicted behaviour and actual manoeuvers of other ships

The actual track and speed of each other ship (TG) and run were compared with the initial expectations of each own ship (OS) on how they will manoeuver. A strong prediction means that the TGs are manoeuvring according to what is expected by the OS (Table 8).

Table 8: Comparison of predicted behaviour and actual manoeuvers of other ships

The expected behaviour of…	ALPHA		BRAVO		CHARLIE		ALL SHIPS	
was correctly predicted	12	56.3%	19	72.2%	17	56.1%	48	61.5%
was moderately correctly predicted	6		7		6		19	
was incorrectly predicted	14	43.7%	10	27.8%	18	43.9%	42	38.5%
Sum	32		36		41		109	

The data indicates that participants are not necessarily very good in guessing what other ships will do, e.g., the expected behaviour of ALPHA and CHARLIE was correctly predicted by the other ships in only a bit more than 50% of the cases. Even though the participants in many cases did not foresee the manoeuver of the TGs, the execution of their own plan was not affected very much (see section 13.4). Also, as strong and moderate association are bunched together which indicates that only minor adaptions were needed to solve the traffic situations even if the TGs did not behave as expected. When considering exclusively the number of strong prediction incidents, it becomes more obvious that the participants needed to adapt their initial plan (albeit not necessarily by much) due to the non-expected behaviour of the other ships which in turn also correlates more with the figures by what AIM suggested.

There are several possible explanations as to why a wrong prediction of the expected behaviour of the other ship(s) does not necessarily mean that the original plan needs to be changed (significantly):

- The predicted behaviour for a give way vessel may either be a change of course (starboard or port) or change of speed making it a less than 50% chance of predicting the manoeuver correctly but neither would necessarily imply a change of manoeuver by the OS.
- The predicted manoeuver for a stand on vessel would according to the COLREGs be to keep the course and speed. However, the stand on vessel may opt (depending on the situation) to take very early action to avoid a close quarters situation by manoeuvring herself. Such manoeuver may not have any influence on the planned and executed action of the OS if such action is planned for another TG.
- Due to navigational constraints, such as nearby land/shallow waters, an OS may expect a TG to execute a manoeuver not strictly according to the COLREGs (albeit at a big distance) whereas from the TG perspective these constraints may not be relevant.
- Although a TG's predicted behaviour may not be correct, a small adaption of the plan would be sufficient to solve the situation, e.g., in an overtaking

situation the overtaking OS may choose to change the course to starboard expecting the TG to keep her course and speed. If this TG is involved in a crossing situation as in the Halland scenario she may also change her course to starboard which may lead the OS to further change the course to starboard but basically still follow her plan.

The analysis of the data also indicate that test persons may not reflect in deeper detail on the possible close quarters situation which may exist or develop between two TGs but mainly focus on the situation of their OS in relation to each TG and choose a manoeuver which solves their situation. This focus needs obviously to be the participant's main concern when solving a traffic situation but may explain the rather poor performance in predicting behaviour of other ships without necessarily having an impact on the OS planned manoeuver.

13.6 Possible AIM influence

The potential influence of AIM is based on the participants initial plan and actual execution (see Table 9). The primary suggestion, and all alternatives were reviewed and compared. In cases in which it was difficult to determine the influence of AIM based on actions, a review of the AIM station's video recording of the complete simulation run was performed to examine how the participants interacted with the suggestions from AIM (i.e., trying to click "accept maneuver" or checking the suggestions often). This helped confirm if the participants may have been influenced by any of the AIM suggestions. The potential outcomes could be:

- Likely: the execution was according to AIM but not according to initial plan
- Possible or inconclusive: The initial plan was either as AIM suggested or the participant seemed to adapt the initial plan as per AIM suggestion, e.g., speed reduction precisely as per AIM
- No: the participant executed the intended plan which was not supported by any AIM suggestions

Table 9: Possible AIM influence

	AIM influence			
	ALPHA	**BRAVO**	**CHARLIE**	**ALL SHIPS**
Likely	4	6	1	11
Possible or inconclusive	1	0	1	2
no	6	5	5	16

In the case of CHARLIE: pure stand on vessel in Anholt and Fehmarn scenarios (i.e., to keep course and speed), was somewhat confusing for some participants as AIM did suggest a manoeuver anyway but the suggested manoeuver needs to be regarded as "last resort" manoeuver in case the other ships don't keep clear.

13.7 Human factors result from thematic analysis

13.7.1 Cognitive Implications

The navigators described their experiences of interacting with AIM in many ways. However, the primary commonality within their descriptions was that they perceived AIM as a tool to which they could evaluate their mental model against, to either confirm or challenge their current situational understanding and awareness. This resulted in both positive and negative reflections from the participants. When suggestions matched their mental model, they responded generally positively towards AIM, as it confirmed their plan of action. Sometimes human characteristics or metaphors were used to get their point across. The following quotes are selected to describe this result.

"It was a confirmation of my decision which felt good, I had my "buddy" here".

"I got the feeling for one moment, that I have a co-pilot. With a computer, that was very good. I don't always listen to the co-pilot. Some minutes I felt some role in the pilot/co-pilot thinking. I appreciate it very much."

Other phrases used to describe AIM include *"consultant", "option generator"*, "*confirmation tool*", and "*thought checker*". However, many times AIM suggestions did not align with the participants plan which caused mixed reactions. Generally, if this mismatch occurred early in the scenario, the participants strongly criticized the suggestions and quickly disregarded AIM as a useful tool and even classified it as potentially dangerous. The participants indicated that many of the risks related to having inaccurate or conflicting suggestions depend on the experience and knowledge of the user. These risks relate to misuse and over-reliance in this technology. However, if suggestions aligned with their mental model earlier on, it seemed that they were more tolerant of mismatched suggestions later in the scenario. In some cases, it was also appreciated to challenge their mental model, as they were still responsible for choosing and implementing the correct course of action. This allowed the participants to consider alternative suggestions that they normally might not have thought about. The following passage summarizes this line of thinking:

"*Maybe it can give me a suggestion I didn't think of. It's also okay if it doesn't' match what I planned but it has to follow COLREG, anti-grounding, anti-collision, etc. Of course, you get more confident if you plan to do something and then it confirms what you were planning to do.*"

13.7.2 Visualisation tool

Another element of AIM which contributed to the navigator's situational awareness and mental model development was the ability to visualize a traffic sequence. This was primarily provided by the "play ahead" function in which the participants could select a suggested manoeuver, and the program would play out how the OS ship would end up if following the suggestion provided the TGs keeping their present course and speed (for an example see Figure 12). This feature was frequently compared to the existing ARPA trial manoeuver and was coined as the "advanced trial manoeuver". Throughout the analysis it became evident that for the navigators, this was one of the biggest contributions of AIM for improving the safety of navigation. This functionality serves a practical use through the ability to visualize the possible consequences of a manoeuver and an estimation of when they should go back to their original route. These aspects of navigating appear to be difficult to do with and without existing navigational aids, even amongst some of the most experienced navigators. This was captured in the following passage:

"*What you see when you look at the radar is the present situation. I start to evaluate myself, of course, I cannot do as much as the computer can, putting in all the facts in the computer. For example, play ahead. Let the computer do the mathematical work, and then I try to use my seamanship to make a decision and evaluate everything*"

Another aspect of visualization that AIM contributed towards was "*outside the box thinking*". This related to the ability to visualize potential emergency manoeuvers if needed, which was both appreciated and potentially misunderstood by some participants. In certain cases, participants thought that AIM was suggesting a 360 degree turn as an equally valid manoeuver since it is not labelled as an emergency manoeuver. This was discussed in most of the interviews and was noted as a misunderstanding. However, generally it was described more positively which is captured in the following passage:

"*Most important, this tool gave a visual suggestion which is good, more input, good alternatives that I hadn't thought about it. It even showed me that I could make a full turn, in case everything goes badly, it can help me get outside the box thinking.*"

13.7.3 Critical elements of collision avoidance decision support systems

Throughout the analysis it was determined that although AIM successfully filled a gap in terms of a navigational aid, many more gaps were identified. Several themes emerged which reflect critical elements needed to further develop automation to ensure that navigators are supported.

13.7.3.1 Integration of systems

Participants identified that continuously adding on new technology without integrating existing systems is a major issue, a practice which is well-known in the maritime industry. This was described by a less experienced navigator in the following way:

"*My main problem with this is that it is not integrated. I am fairly against adding systems and screens. You can call it a tool, but it is not a tool if you are just adding it like a secondary system. I won't focus on it. Looking for ECDIS to be a more integrated system. Adding functions in ECDIS won't help me. You have more functions than most people know about. It is unnecessary, it's complicated, and it's hard to find. I am new, so I rely on Radar and AIS for draft info, etc. but nothing else. I feel like we need to make ECDIS more integrated system from the beginning before adding more tools*."

13.7.3.2 Automation transparency

Automation transparency emerged as a theme in two parts; (1) what participants want out of a decision support system, and (2) how information could be shared with other vessels to create a shared awareness in traffic situations. The first aspect of transparency relates to the functionality of a decision support system. Participants identified that to develop trust in a technology, there must be a proper foundation and training regarding the systems opacity, with a clear understanding of the capabilities, and potential risks of the system. This would lead to better human-automation interaction, where an appropriate level of reliance can be placed upon the technology. This was described in the following way:

"*If you add a tool like this, it shouldn't be added to the bridge without proper training for the officers. If you have a good training, you will learn what advantages you can take with this tool and what risks it might add. It is a prediction tool that you can use in many steps ahead. 8/10 times you can rely on it but must always use good seamanship and use your own skills. You also must not be afraid to say that AIM is wrong. Must make sure it is your own decision. Same with ECDIS incidents, ARPA, etc. Same discussion all the time, comes back to overreliance on systems and complacency. It is so easy to trust the systems for the human mind when you see it on the screen than reality. The mind finds it easier to accept that*."

Participants also explored the notion that ships should "*talk to each other*" which reflects the concepts of information and route exchange explored within the STM project (see introduction for applicable references). There were suggestions to integrate AIM with a route sharing function which they believe could provide a more complete situational awareness and potentially reduce the number of assumptions related to how other vessels might manoeuver. The navigators want technology that supports them. Although already very good at their job, they remain open to trusting and accepting technology that allows for more informed decision making. Most participants are hopeful for the possibilities of automation yet maintain considerable hesitancy. We believe this hesitancy stems from their experiences with existing technologies onboard, and the accompanying challenges. Increasing transparency in automated systems would improve the human-automation relationship so that a practical trust and reliance in the technology can be developed through an informed understanding of its abilities and limitations.

13.7.3.3 Blunt but useable system

This category merged two seemingly contradicting themes, (1) AIM is a blunt tool, and (2) AIM was very user friendly. It was evident amongst the participants that AIM was limited in what it could do, resulting in the blunt description. AIM was effective from OS perspective in that it had an "egocentric" perspective but lacked the birds-eye overview of the entire traffic situation to consider situations between other ships. This was seemingly something the participants hoped of this technology, which AIM could not provide. It was discussed that AIM is simple because its suggestions are based on mathematical calculations of heading vectors. The task of navigation is complex, and possible through so many different aspects of knowledge reaching far beyond the capacity of AIM. This can be described by the following passage:

"*In the hands of a good navigator it is a supportive tool but not more than that. The algorithm clearly states that it is to follow the collision regulations and that it is a decision support tool. Many suggestions that you get from the algorithm will not be used in real life. You could drive yourself into a situation that you cannot handle. Most experienced navigators think about what could/will happen if I end up in a situation/something happens. If engine stops, etc. From collision regulation avoidance it can be good supportive tool but beyond that I think not*."

Yet, surprisingly, the participants still really liked AIM. This caused slight confusion amongst the research team for some time as we struggled to reason with why they liked this tool so much if it was so simple? Part of the answer emerged when we analyzed the comments related to the user interface and overall usability of the system which resulted in an overwhelming positive response. Unfortunately, user-centered solutions are few and far between in the maritime industry. Therefore, one of the biggest successes of AIM is that it is easy to use and does what it is supposed to do, even if it is blunt. This was described simply by one of the participants, and reiterated by almost everyone:

"*Extremely easy to use. Interface was easy to use. No unnecessary functions and it did what it was supposed to do. It was straight forward*."

13.7.4 Implications for seafarers

13.7.4.1 Seamanship – "a floating abstract norm"

The navigators identified that seamanship and the formal regulations are heavily intertwined, as described in the quote below. Untangling the role of seamanship within new technology is challenging because even the COLREGs mention that action shall be taken "with due regard to the observance of good seamanship" (Rule 8 of COLREGs). There were discussions about what it is, if an algorithm could have it, if the AIM suggestions portrayed good or bad seamanship, and how important it is in the decision-making paradigm. Given the fact that all navigators claim to know what good seamanship is, yet it is defined and identified differently by everyone, these questions proved difficult to answer. However, a common factor was that the navigators perceive seamanship as having a good overview of the situation, including an assessment of "*how my actions affect other vessels in the area*", something which AIM is not programmed to do. The participants unanimously agreed that AIM does not have "seamanship", nor is it possible for an algorithm to have seamanship. However, they understood that the settings could be adapted to receive suggestions at different times and distances from other vessels.

"*If you go against good seamanship, you go against the COLREG – the only law text in the world that is included in COLREG*"

13.7.4.2 Potential deskilling

Participants were conflicted about the implications of having manoeuvring suggestions which might impact motivation and hone the skills to safely assess the route based on all available information. Some participants thought that this could cause increased complacency amongst certain navigators who might take the "easy way out" without properly thinking for themselves, especially less skilled or novice navigators. They also identified that these risks exist with existing navigational aids, including ECDIS and radar. However, with these risks in mind, most participants argued that knowledge of the COLREGs might be even more important when using AIM or similar tools and the core knowledge of navigation in education should not be reduced because of this type of technology. This was described by a participant in the passage below.

"*If you take it a few steps more, I am worried that some things will disappear that are vital in the "art of navigation". The craftmanship should be more important to get from the school, it is very important that it doesn't get replaced with courses for AI/technology. We have 3 years of studies, and time is limited. Basic skills must still be there and practiced, from my point of view.*"

This passage also conveys an interesting point about the description of navigation as an "art". AIM assumes that other ships will keep their course and speed which provides the user with a mathematical solution to one aspect of a complex problem. To solve this complex problem, these mathematical solutions must be cross-checked by the navigators. The ability to do this requires an understanding of the entire navigational situation, which is possible through the many years of formal education involved in training navigators to solve navigational situations as more than mathematical equations and more like an "art".

14 Addressing the Research Questions

Decision support is the next step in the transformation towards more connected smart ships, however, such systems for navigation are largely unexplored in existing research from a human perspective. This research provided insight into some of the potential impacts associated with the integration of a decision support system in basic traffic scenarios. Four research questions were answered throughout this project:

- Can experience, good seamanship, situational awareness and other non-technical skills that impact upon safe and efficient navigation be implemented into algorithms?
- Can traffic situations that have both human operators and smart/autonomous vessels be safely resolved compared to situations encountered today?
- From the navigator's perspective, what is the role of a collision avoidance decision support system for navigation?
- From the navigator's perspective, what are the benefits and potential challenges associated with a decision support collision avoidance system?

14.1 Can experience, good seamanship, situational awareness and other non-technical skills that impact upon safe and efficient navigation be implemented into algorithms?

The COLREGs would seem to be the obvious departure point for developing AI and automation solutions. Considering that COLREG Rule 8 states that any action taken to avoid collisions shall, if the circumstances of the case admit, be positive, made in ample time and with due regard to the observance of good seamanship, AIM does provide some means to actively set values which may reflect the operator's interpretation of at least "positive" and "ample time". However certain decision-making characteristics which have evolved in navigation practices may create difficulties for those on the technical side of AI to use appropriate machine/deep learning approaches to develop situation-appropriate AI. While the wording of the COLREGs is sufficiently precise, actions depend on a navigator's ability to use common sense ("good seamanship") to not only determine if a situation currently applies, but also to exploit the flexibility in the actions prescribed in a rule (MacKinnon, Weber et al. 2020).

As an example, AIM's primary suggestion on Alpha in the Halland scenario was a course alteration to port with Charlie being the primary target (Figure 24). Bravo's status was given as a head-on situation meaning that both Alpha and Bravo need to take action and according to Rule 14 change their course to starboard which was the AIM suggestion on Bravo (Figure 25). If both ships follow the primary suggestion, the manoeuver may potentially result in in a new but different COLREG situation i.e., a crossing situation with unknown further consequences.

Figure 24 Primary suggestion on Alpha in Halland scenario

Figure 25 Primary suggestion on Bravo in Halland scenario

The example in Figures 24 and 25 may illustrate the complexity of a seemingly straightforward traffic situation involving 3 ships and the challenges of coding an algorithm:

1) The primary suggestion on Alpha is based on identifying Charlie as the primary target probably due to her being closer in range. However, all 3 ships in the scenario have virtually the same TCPA to one another. This raises the question whether the identification of a single primary target based on range is sufficient.

2) The suggestion on Alpha may be against Rule 14 but considering that Bravo is at a distance of 12 NM, it may be argued that, although the ships are in sight of one another, the COLREGS don't apply at such distances. However, distances alone are not sufficient to judge the application of the rules as any action according to Rule 8 needs to be taken in "ample time" which is obviously based on the speed of the ships involved in the traffic situation.

3) As Alpha is constrained by land on her starboard side and bearing in mind that Bravo is at a safe distance, a distinct course change to port to overtake Charlie may be considered as "good seamanship" or in other words a safer action than overtaking Charlie on her starboard side as proposed in AIM's alternative suggestion (Figure 26). This is even more relevant when expecting Charlie to alter course to starboard to avoid a close quarters situation with Bravo.

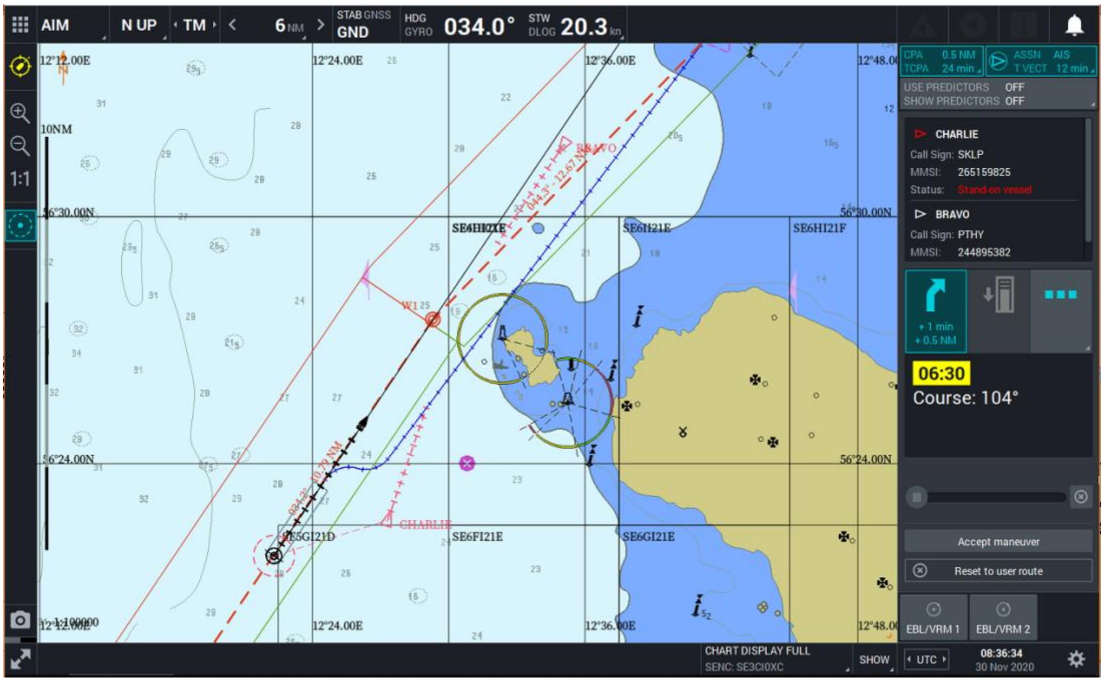


Figure 26 Alternative suggestion on Alpha in Halland scenario

4) The possibility of Alpha to reduce speed and stay astern of Charlie (including possibly a minor course alteration to starboard) and expecting Bravo to sufficiently alter course to starboard avoiding both Alpha and Charlie may also be considered although was not a suggestion by AIM.

5) Finally, Alpha may assume Bravo, having due regard to the constraints of Alpha and Charlie, having land on their starboard side will sufficiently alter course to starboard allowing Alpha to overtake Charlie on her port side with a safe CPA.

Based on the track records of the simulations there is an indication that the majority of the participant's acted according to option 5 (Figure 27 and 28). Red and yellow tracks indicate runs without AIM (red) and with AIM (yellow) respectively. The thin white line is the planned OS route and the thin green and red lines AIM suggestions.

Figure 27 Tracks of Alpha in Halland scenario

Figure 28 Tracks of Bravo in Halland scenario

The data are not statistically significant due to the limited number of trials, but it raises the question whether option 5 (which may be regarded as a violation of the COLREGS) can or even should be programmed in an algorithm. It also needs to be remembered that the suggestions provided by AIM were based on certain CPA/TCPA values being identical on all involved ships and different pre-selected settings will have resulted in different solutions. Regardless, humans may rely and act on potentially well-founded assumptions that other ships act according to the poorly operationalized and possibly area dependent term "good seamanship", algorithms will likely need significant data based on deep machine learning which is difficult to obtain. However, even if such data may eventually be available, the fundamental question of what is considered as possibly violating the COLREGS versus acting according to "good seamanship" remains.

Also, humans are flexible to alter and adapt their plan to a lesser degree which is reflected in Table 7 where participants in 88% of the cases were able to follow their original plan to a large or moderate extent. The interesting number there is that in 31% of the cases a minor or moderate adaption was needed to avoid a close quarters situation. If such flexibility and moderate adaption of a plan constitutes a major factor in avoiding close quarters situation cannot be answered by this study nor whether algorithms may or should be trained or not to incorporate such flexibility.

14.2 Can traffic situations that have both human operators and smart/autonomous vessels be safely resolved compared to situations encountered today?

This research question stems from the original plan to have autonomous ships interact within the same scenarios as human-operated ships. As the project evolved, the researchers discussed the relevance of the data/variables and it was decided that the human element was the priority of the main data collections, as this presents the largest existing research gap. Additionally, autonomous trials would reduce the ability to gain input from human operators while still using simulator time. This was further motivated by the fact that currently AIM is developed as a decision support system, and not as an autonomous system. Therefore, fully autonomous vessels trials were not included in this study in order to prioritize data from the human-operated trials. This research question remains very relevant and is suggested as a next step for future research.

Throughout the trials there are certain data which can provide insight into this research question. The primary suggestion from AIM would be the manoeuvre that would be applied in "autonomous mode". This provided insight into what the autonomous vessel would be programmed to do and allowed for a comparison with what participants predicted the vessel would do. The data comparing predicted behaviour and AIM's default/primary suggestion on the TG was relatively low at about 50%, meaning that participants expected a different behaviour in 50% of the cases (Table 6). However, this number must be interpreted with the data comparing expected behaviour and actual manoeuvre, which was only slightly better, i.e., about 60% of the participants correctly predicted the manoeuvre of the other ships in the scenarios (Table 8). As the COLREGs do not stipulate a specific manoeuvre to avoid a collision and that collisions from e.g., a give way ship perspective may be avoided by either a change of course or speed, the prediction of how a ship may behave in a traffic situation even if a low number of ships are involved, is not straightforward. Regardless, in almost 90% of the cases, participants followed their initial plan with mostly minor adaptions which indicates that although a participant expects a different manoeuvre of a TG ship, such manoeuvre may not give rise to a complete change of plan on the OS (Table 7). Therefore, the assumption that the risk of a close quarter situation is reduced if a ship acts as expected by another ship may not necessarily be fully correct nor may an incorrect prediction have a major influence on a planned manoeuvre. Experienced navigators commonly make small adaptions of course and speed in a timely manner to avoid upcoming close quarters situations or adapt an ongoing avoiding manoeuvre. It is difficult to envision how this could be implemented by algorithms, particularly for traffic within confined waterways. Further research is needed to understand these nuances as more advanced algorithm-based systems are developed for collision avoidance.

Another factor to consider related to mixed human and automation traffic situations is communication. Results from the Sea Traffic Management (STM) Project found that communication between ships and between ship and shore was one of the most important factors to consider as various types of automation are added to existing work practices (Aylward 2020, Aylward, Johannesson et al. 2020, Aylward, Weber et al. 2020). Introducing any automated function/tool which shares information between ships, or suggests routes based on the most likely actions of another ship, is providing an additional means of indirect or nonverbal communication that is not available today. In this study, VHF ship to ship communication was available. However, it was almost never used to solve any traffic situations. While this is fully in accordance with various guidance notes in the maritime industry which highlight the danger of attempting to avoid close quarters situations by using VHF radio, it may not reflect the work as done in current navigation practices. Today, there remain a significant number of collisions where subsequent investigations have found that at some stage before impact, one or both parties were using VHF radio in an attempt to avoid collision (MCA 2017). Additionally, all existing research and data surrounding maritime traffic communication is based on VHF radio interactions as this is the accepted and legal means of communication today. Implementing additional ways to communicate directly and indirectly have the possibility to have unintended effects without systematic and vigorous verification and validation of these practices (Aylward, Johannesson et al. 2020). Finally, considering the implementation of autonomous vessels in traffic situations with manned ships the uncertainty is even greater, and more research is needed to specifically address maritime traffic communication.

14.3 From the navigator's perspective, what is the role of a collision avoidance decision support system for navigation?

Part of understanding the role of a new technology is determining the need for it through a rigorous human-centred design approach. To date, decision support systems have largely ignored the capacities and needs of the end-users. In the early stages of technology testing, it is important to develop scenarios which fit within the technology operational limits but can also provide valuable information from the participant perspective. Given the dynamic nature of navigation, it was crucial to provide the participants with different opportunities and traffic situations (crossing, meeting, and overtaking) to assess the technology. AIM is intended for use in open sea, and the scenarios tested in this study potentially pushed its operational limits (see scenario descriptions in Table 1). When discussing the optimal use case for a decision support system like AIM the participants responded that it would be less useful in complex or busy traffic situations and that they prefer to rely on other information sources (i.e., looking out the window and ARPA) to make decisions, aligning with previous works (Hochgeschurz, Motz et al. 2021). Therefore, according to the participants it seems that given their high workload in busy traffic situations, they would rely on it more in less busy traffic situations to prepare and plan. This might seem counterintuitive yet could be attributed to the fact that humans are flexible and can adapt to unexpected situations (Hancock, Jagacinski et al. 2013, Janssen, Donker et al. 2019). The safe navigation of a ship requires flexibility and adaptability due to the inherent complexity of traffic situations at sea.

From an information processing perspective, digitalised information is intended to support human in the process of decision-making, providing the most useful information to the operator (Parasuraman, Sheridan et al. 2000). Parasuraman et al. proposed four levels of input functions associated with human information processing stages, Level 1: information acquisition, Level 2: information analysis, Level 3: Decision selection, Level 4: Action implementation (Parasuraman, Sheridan et al. 2000). The AIM system can provide automated support for Levels 1-3 of input functions. AIM provides information acquisition (Level 1) through supporting the human sensory process of organizing incoming data from surrounding ships, information analysis (Level 2) through mathematically predicting which route is the most optimal and predicting when the user should proceed back to the route, and partially decision selection (Level 3), although there is no replacement of human decision making there is augmentation, as the system provides several different alternatives of potential manoeuvres to solve a traffic situation. The results indicate that AIM provided the ability to confirm or challenge a navigator's plan of action through visualizing how the traffic situations could be solved according to the COLREGs. This ability to visualize in this way is currently not available through other navigational aids, thereby supporting the human sensory, memory, and inferential processes. This caused mixed reactions for participants depending on their overall experience interacting with AIM and how they perceived it.

The results show that at the beginning of the scenario, the participants and AIM approached target ships in a similar way, 79% of the time identifying the same primary target (Table 4) and planning the same manoeuvre 72% of the time (Table 5). These data describe that at the beginning of the scenario, their decision was confirmed most of the time. This also suggests that the participants approached their navigation plan from a primarily rule-based perspective grounded in the COLREGs, which is the basis of the AIM algorithms. Maritime accident data in recent years indicates that 56% of collisions at sea are caused by violations to the COLREGs (Statheros, Howells et al. 2008, Liu, Zhang et al. 2016, Abilio Ramos, Utne et al. 2019). Live traffic situations are complex and dynamic which are safely resolved based on several factors including, adherence to the formal rules (COLREGs), the informal rules, and their interpretation of good seamanship. Therefore, it seems that AIM was able to support the operator primarily in one aspect of their decision making. AIM allowed the navigators to check if their plan agrees or conflict with the rules and visualize (play-ahead) how certain actions would unfold. Therefore, although AIM was described as a blunt tool which primarily contributed to the mathematical calculations or strict application of the COLREGs, the participants believed that even its basic functionality has an important role in the safety of navigation.

14.3.1 Advisory system

According to the participants, the primary benefits of this decision support system was to be able to visualize a manoeuvre of a potential future traffic situation based on suggestions generated by the system. This feature called "play-ahead" can contribute to a more complete overview of a situation and the ways it could unfold, while keeping in mind that this function is based on the TGs keeping their course and speed (which may not always be a correct assumption). Other safety-critical industries are continuously working to utilize automation to support operator performance. A computer operator support system (COSS), also known as a "recommender system" or "operator advisory system" is a system with a long history in the nuclear power and aviation industries that can assist operators in monitoring performance, diagnosing faults, predicting future states, recommending mitigations, and decision support (Boring, Thomas et al. 2015, Westin, Borst et al. 2016). The basis for an operator advisory system is that it supports the operator with a task and aids them in the completion of the task when possible while being minimally intrusive to manual operations (Boring, Thomas et al. 2015). Although AIM is a decision support system, perhaps it should be sub-classified as an advisory system given the broad scope of decision support systems. Advisory systems are specifically used to support the human decision-making process in unstructured, complex, or open-ended situations, such as navigation. The participants even described AIM as an "option generator", "buddy" or "co-pilot" aligning closely with the synonyms presented for such systems. ¨Shipping 4.0¨ will provide an abundance of automated systems and features and there should be a better, more descriptive classification of what type of decision support is provided to the operator. The maritime industry should adopt lessons learned from other safety critical industries which have successfully adopted and integrated similar systems.

14.4 From the navigator's perspective, what are the benefits and potential challenges associated with a decision support collision avoidance system?

The research question aimed to understand what the potential benefits and challenges are associated with using a decision support collision avoidance system. AIM is one technology of many that could provide decision support to navigators. Therefore, the focus of this study was not to dissect the individual suggestions of AIM throughout the entire scenario, but rather to explore the nature and potential impacts of using a decision support system for navigation as they are the next step towards the implementation of the MASS concept (IMO 2021). The human will remain in control and can choose how to use the technology and the ship will function as normal. However, if navigators opt to use a decision support system, there will be an impact on numerous aspects of navigation as it exists within this complex sociotechnical system. Small changes in the system architecture can transform judgements, roles, relationship and weightings on different goals resulting in vast changes of system function (Woods and Dekker 2000, Grech, Horberry et al. 2008). Potential impacts are discussed from a human-automaton interaction perspective and the role of the human element in maritime systems.

14.4.1 Human-automation interaction

Human-automation interactions are complex and the trust and reliance in automation evolve based on personal history, cultural and organisational factors may not have been traditionally present in some highly automated industries (Lee and See 2004). In this study the participants tested a technology for the first time, meaning it was difficult to evaluate their level of trust in the system. Throughout the interviews several themes emerged which provided indications of how trust could be developed in this technology. At the advisory level of a decision support system, it is important to have transparency in the automation to develop the human-automation relationship. The navigators wanted to know more about how and why suggestions were being presented. Communication to the user about how, what, and why, information is being presented, also known as system opacity, is related to trust and acceptance of automation (Westin, Borst et al. 2016). With more transparency, the navigators can better evaluate whether they want to use this information instead of trying to evaluate why the suggestion was presented. A balance must also be achieved between automation transparency and information overload. Further work is needed to determine the critical level and type of information needed for different aspects of navigational tasks.

Predicting the behaviour of another ship and correctly understanding its intentions is crucial to safe navigation. Uncertainties, in particular the inability to anticipate the actions of another ship have been listed as a causal factor in most collisions between 1999 and 2012 (Langard, Morel et al. 2015, Wickens, Williams et al. 2020). The results from this study revealed that the participants performed relatively poorly in predicting the actions of other ships, considering the scenarios only involve three ships (Table 8). The inaccurate predictions never led to any unsafe situations as the participants were easily able to adapt to the traffic situations, most participants (88%) were even able to maintain their original plan (Table 7). However, it is an important finding to discuss as it reveals a weakness in the operator's situation awareness, particularly in the projection of future states of events (Endsley 1995). This also highlights an aspect of human information processing which could benefit from the use of automation and should be considered for future developments of decision support systems. It is critical that automation supports humans in their weaknesses (e.g., visualization and computation), allowing them to use their strengths (seamanship) to have a more accurate situation awareness.

14.4.2 The regulatory element

One of the challenges identified by the participants was related to the lack of integrated systems, a long-standing problem in the maritime industry (IMO 2006). It is important to note that ECDIS manufacturing is strictly regulated (prescriptive standards) which does not allow the manufacturers to freely implement new ideas, design, or features. Wärtsilä has solved the issue by implementing AIM as a supplementary tab/menu in their Beta version of an upcoming ECDIS software suite, thereby avoiding the restrictions imposed by the performance standard. In this study AIM was presented to the users on a separate screen, however, in practice AIM will be semi-integrated. Although highlighted as a priority for IMO (IMO 2006), stand-alone, distributed systems remain a significant hurdle in the development and implementation of navigation systems. AIM was described as helpful and useful, yet navigators remain hesitant to accept any new systems without addressing the existing distributed systems on the bridge, before adding more. One approach to address the integration of systems and standardization is through the OpenBridge project (Nordby, Gernez et al. 2019, Nordby, Gernez et al. 2020). This project is proving successful amongst major maritime stakeholders and would contribute to safer maritime user interfaces. This challenge must be addressed by IMO at the regulatory level and should be prioritized moving towards more digitalized solutions.

Another challenge identified by the participants was the potential impact of decision support systems on seafarer training, skill maintenance and development. The role and qualifications of the future seafarer has been identified by IMO as one of the most complex issues to be addressed with the adoption of MASS (IMO 2021). The future competencies of seafarers and their role within the maritime ecosystem has also been explored by several academics (Man, Lundh et al. 2015, Baldauf, Fischer et al. 2019, Mallam, Nazir et al. 2019, Kim and Mallam 2020). However, most of these works target higher levels and degrees of automation and MASS adoption, while little is known about the role of the seafarer for low levels of MASS (i.e., Degree 1). According to observations from this study, the potential impacts should not be ignored and there is a possibility for skill degradation without proper training and formal introduction to automated systems. The participants almost unanimously agreed that while seafarers are still on board and in control, the education, training, and "core navigational knowledge" is essential.

It was further identified that the potential dangers associated with the use of any automated system including complacency and over-reliance should be taken seriously. These risks are also present with existing navigational aids, including ECDIS and radar which was clearly noted in IMO MSC 82/15/2 (IMO 2006). The participants were clear that the technology manufacturers should not market these systems towards inexperienced, fatigued or poorly educated officers. Instead, at this early stage of MASS adoption, decision support should be advisory in nature and provide well trained officers' rule-based information (COLREG) to make a final decision for safe navigation.

15 Conclusions and future work

This project explored the use of a collision avoidance decision support system for navigation. The results provided insight into how the AIM decision support system was perceived by navigators and how it compared to human decision making. AIM played an important role in allowing the navigators to check their navigational decisions which were either confirmed or challenged by AIM. This is possible through a visualization of potential future traffic situations and was identified as an important addition to existing navigational aids and a complement to human information processing. AIM was appreciated by the users and most participants indicated they would like to use it should it become available to the bridge technology complement. However, the integration of any technology can have wider impacts on the maritime sociotechnical system and AIM is no exception. There are concerns with trusting and relying on the information from AIM, which could be improved with better user-centered design, greater system transparency and additional user testing. Certain findings, primarily the integration of systems and the training and skill development of today and tomorrow's seafarers must be addressed by regulatory bodies. The transition towards MASS is exciting, digital solutions will change the shipping industry as we know it. However, we must proceed with caution and attempt to fill some of the major research gaps facing today's seafarers to better prepare for tomorrow.

A quick literature search will reveal hundreds of research efforts focused on remote and autonomous systems looking far into the future towards an undetermined timeline of exclusively MASS Degree 3 and 4. Although this research may be critical to achieve safer and more efficient future maritime systems, there is a very slow adoption of remote and autonomous shipping in the industry. A deeper dive into the literature reveals a major research gap focusing on the description of current and future maritime sociotechnical systems. Only a handful of studies have been completed to understand better MASS Degree 1 and 2 which focus on human interaction with decision support systems and mixed systems of varying levels and types of automation. Therefore, there are many constructs which should be further explored:

- In the near-term end-users are involved in all stages of the development of automated functions to ensure they complement the existing working environment and human capacities. As the ship becomes more complex, the role and work practices on board will change. Therefore, it is suggested that there is an increased focus on systematic testing and integration of automated functions with continuous user testing and feedback.

- Human-automation interaction studies focusing on specific human factors including trust in automation, use and acceptance of automation, reliance of automation, situation awareness, workload, work tasks, training, communication, and the role of the future seafarer.
- Mixed traffic scenarios of human-operated and autonomous ships. Results from this project compare and contrast human decision making and algorithm-generated solutions to traffic situations. This interaction should be better understood as this will represent a realistic traffic environment in the coming years.
- The adoption of an industry wide holistic user-driven framework for the development, testing, and implementation of new technologies which allow for more systematic testing and validation activities.

Regarding the quality of the algorithm used and especially focusing on the identification of the primary target and the default suggestion presented by AIM, it matches the participants own plan to a large extent (79% for primary target and 72% for suggestion). Considering that the COLREGs do not necessarily specify a single action to avoid a collision, these figures suggests that the AIM algorithm reflects to a rather large extent how the participants approached their navigation plan from a primarily rule-based perspective grounded in the COLREGs. Additionally, by suggesting alternatives, AIM does in principle cover most possible actions even though some of them were discarded as not practicable or even incorrect in certain cases. However, for the transition to fully autonomous ships there are several issues concerning the algorithms which need to be further explored:

- The study only covered traffic situations where the CPA/TCPA settings in AIM were equal on all ships. However, these settings are user defined and will need to be variable not only due to traffic density, geographical and navigational limitations but may also be different depending on a ship's manoeuvring characteristics. Developers of algorithms will need to take these parameters into account and regardless of implementation there may be traffic situations involving autonomous ships with different CPA/TCPA settings and therefore unknown consequences.
- The study showed that in 31% of the cases, participants made a moderate adaption to the original plan to avoid a close quarters situation. If such flexibility and moderate adaption of a plan constitutes a major factor in avoiding close quarters situation was not answered by this study nor whether algorithms could or should be trained or not to incorporate such flexibility and minor adaption capabilities. Further studies involving more complex traffic situations are required.
- Humans may rely and act on potentially well-founded assumptions that other ships act according to the poorly operationalized and possibly area

dependent term "good seamanship". Algorithms will likely need significant data based on deep machine learning which is difficult to obtain. However, even if such data may eventually be available, the fundamental question of what is considered as possibly violating the COLREGS versus acting according to "good seamanship" remains.

- This study only investigated traffic situations involving three ships where all ships met after approximately the same TCPA and similar CPA. To further test and exploring the effect of algorithms, multi ship including sequential encounters in traffic situations need to be simulated.

16 References

Abilio Ramos, M., I. B. Utne and A. Mosleh (2019). "Collision avoidance on maritime autonomous surface ships: Operators' tasks and human failure events." Safety Science **116**: 33-44.

Aylward, K. (2020). Automated Functions: Their Potential for Impact Upon Maritime Sociotechnial Systems Licentiate Thesis Chalmers University of Technology

Aylward, K., A. Johannesson, R. Weber, S. N. MacKinnon and M. Lundh (2020). "An evaluation of low-level automation navigation functions upon vessel traffic services work practices." WMU Journal of Maritime Affairs.

Aylward, K., R. Weber, Y. Man, M. Lundh and S. N. MacKinnon (2020). "“Are You Planning to Follow Your Route?” The Effect of Route Exchange on Decision Making, Trust, and Safety." Journal of Marine Science and Engineering **8**(4): 280.

Bainbridge, L. (1983). "Ironies of automation." Automatica **19**(6): 775-779.

Baldauf, M., S. Fischer, M. Kitada, R. Mehdi, M. A. Al-Quhali and M. Fiorini (2019). "Merging Conventionally Navigating Ships and MASS - Merging VTS, FOC and SCC?" TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation **13**(3): 495-501.

Boring, R. L., K. D. Thomas, T. A. Ulrich and R. T. Lew (2015). "Computerized Operator Support Systems to Aid Decision Making in Nuclear Power Plants." Procedia Manufacturing **3**: 5261-5268.

Braun, V. and V. Clarke (2006). "Using thematic analysis in psychology." Qualitative Research in Psychology **3**(2): 77-101.

Creswell, J. W. and V. L. P. Clark (2017). Designing and conducting mixed methods research, Sage publications.

Endsley, M. R. (1995). "Toward a Theory of Situation Awareness in Dynamic Systems." Human Factors **37**(1): 32-64.

Endsley, M. R. (2017). "From Here to Autonomy:Lessons Learned From Human–Automation Research." Human Factors **59**(1): 5-27.

Etikan, I., S. A. Musa and R. S. Alkassim (2016). "Comparison of convenience sampling and purposive sampling." American journal of theoretical and applied statistics **5**(1): 1-4.

Grech, M., T. Horberry and T. Koester (2008). Human factors in the maritime domain, CRC press.

Hancock, P. A., R. J. Jagacinski, R. Parasuraman, C. D. Wickens, G. F. Wilson and D. B. Kaber (2013). "Human-automation interaction research: Past, present, and future." Ergonomics in Design **21**(2): 9-14.

Hochgeschurz, S., F. Motz, R. Grundmann, S. Kretzer and L. Thiele (2021). "Which Radar and ECDIS Functionalities Do Nautical Officers Really Need in Certain Navigational Situations?" TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation **15**(1): 73-81.

Hörteborn, A., J. W. Ringsberg, M. Svanberg and H. Holm (2019). "A Revisit of the Definition of the Ship Domain based on AIS Analysis." Journal of Navigation **72**(3): 777-794.

IMO (1972). Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs). http://www.imo.org/en/About/conventions/listofconventions/pages/colreg.aspx, International Maritime Organization

IMO (2006). MSC 82/15/2 Role of the human element. M. S. Committee. London, Internationa Maritime Organisation, IMO.

IMO. (2021, 25 May 2021). "Autonomous ships: regulatory scoping exercise completed." Retrieved 11/04, 2021, from https://www.imo.org/en/MediaCentre/PressBriefings/pages/MASSRSE2021.aspx.

Janssen, C. P., S. F. Donker, D. P. Brumby and A. L. Kun (2019). "History and future of human-automation interaction." International Journal of Human-Computer Studies **131**: 99-107.

Kim, T.-e. and S. Mallam (2020). "A Delphi-AHP study on STCW leadership competence in the age of autonomous maritime operations." WMU Journal of Maritime Affairs.

Langard, B., G. Morel and C. Chauvin (2015). "Collision risk management in passenger transportation: A study of the conditions for success in a safe shipping company." Psychologie française **60**(2): 111-127.

Lee, J. D. and K. A. See (2004). "Trust in automation: Designing for appropriate reliance." Human factors **46**(1): 50-80.

Liu, Z., Y. Zhang, X. Yu and C. Yuan (2016). "Unmanned surface vehicles: An overview of developments and challenges." Annual Reviews in Control **41**: 71-93.

MacKinnon, S. N., R. Weber, F. Olindersson and M. Lundh (2020). Artificial Intelligence in Maritime Navigation: A Human Factors Perspective. AFHE, Cham, Springer International Publishing.

Mallam, S. C., S. Nazir and S. K. Renganayagalu (2019). "Rethinking Maritime Education, Training, and Operations in the Digital Era: Applications for

Emerging Immersive Technologies." Journal of Marine Science and Engineering **7**(12): 428.

Man, Y. (2015). Human-Machine Interface Considerations for Design and Testing in Distributed Sociotechnical Systems.

Man, Y., M. Lundh, T. Porathe and S. MacKinnon (2015). "From Desk to Field - Human Factor Issues in Remote Monitoring and Controlling of Autonomous Unmanned Vessels." Procedia Manufacturing **3**: 2674-2681.

MCA (2017). MGN 324 (M+F) Amendment 1 Watchkeeping Safety – VHF Radio & AIS. M. a. C. Agency.

Michaelides, M., M. Lind, L. Green, J. Askvik and Z. Siokouros (2020). Decision Support in Short Sea Shipping. Progress in IS, Springer International Publishing**:** 221-236.

Nautilus, F. (2018). Future proofed? What maritime professionals think about autonomous shipping. London, UK, Nautilus Federation.

Nordby, K., E. Gernez, S. Frydenberg and J. O. Eikenes (2020). Augmenting OpenBridge: An Open User Interface Architecture for Augmented Reality Applications on Ship Bridges.

Nordby, K., E. Gernez and S. Mallam (2019). OpenBridge: Designing for Consistency Across User Interfaces in Multi-Vendor Ship Bridges.

Ożoga, B. and J. Montewka (2018). "Towards a decision support system for maritime navigation on heavily trafficked basins." Ocean Engineering **159**: 88-97.

Parasuraman, R., T. B. Sheridan and C. D. Wickens (2000). "A model for types and levels of human interaction with automation." IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans **30**(3): 286-297.

Patton, M. Q. (2002). "Qualitative interviewing." Qualitative research and evaluation methods **3**: 344-347.

Perera, L. P. and B. Batalden (2019). Possible COLREGs Failures under Digital Helmsman of Autonomous Ships. OCEANS 2019 - Marseille.

Pietrzykowski, Z. and P. Wołejsza (2016). Decision Support System in Marine Navigation, Cham, Springer International Publishing.

Porathe, T. (2021). "Human-Automation Interaction for Autonomous Ships: Decision Support for Remote Operators." TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation **15**(3): 511-515.

Porres, I., S. Azimi, S. e. Lafond, J. Lilius, J. Salokannel and M. Salokorpi (2021). "On the Verification and Validation of AI Navigation Algorithms." arXiv pre-print server.

Ramos, M., I. B. Utne, J. E. Vinnem and A. Mosleh (2018). Accounting for human failure in autonomous ship operations. European Safety and Reliability Conference ESREL 2018, Trondheim, Norway.

Ramos, M. A., C. A. Thieme, I. B. Utne and A. Mosleh (2020). "Human-system concurrent task analysis for maritime autonomous surface ship operation and safety." Reliability Engineering & System Safety **195**: 106697.

SAGE. (2008, 2019/11/11). "The SAGE Encyclopedia of Qualitative Research Methods." from https://methods.sagepub.com/reference/sage-encyc-qualitative-research-methods.

Statheros, T., G. Howells and K. M. Maier (2008). "Autonomous ship collision avoidance navigation concepts, technologies and techniques." The Journal of Navigation **61**(1): 129-142.

Vestre, A., A. Bakdi, E. Vanem and Ø. Engelhardtsen (2021). "AIS-based near-collision database generation and analysis of real collision avoidance manoeuvres." The Journal of Navigation **74**(5): 985-1008.

Westin, C., C. Borst and B. Hilburn (2016). "Automation Transparency and Personalized Decision Support: Air Traffic Controller Interaction with a Resolution Advisory System**This work is co-financed by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E in the SESAR Programme. Opinions expressed in this work reflect the authors' views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein." IFAC-PapersOnLine **49**(19): 201-206.

Wickens, C. D., A. Williams, B. A. Clegg and C. A. P. Smith (2020). "Nautical Collision Avoidance:The Cognitive Challenges of Balancing Safety, Efficiency, and Procedures." Human Factors **62**(8): 1304-1321.

Woerner, K. (2016). Multi-contact protocol-constrained collision avoidance for autonomous marine vehicles, MASSACHUSETTS INSTITUTE OF TECHNOLOGY.

Woerner, K., M. R. Benjamin, M. Novitzky and J. J. Leonard (2019). "Quantifying protocol evaluation for autonomous collision avoidance." Autonomous Robots **43**(4): 967-991.

Woerner, K. L., M. R. Benjamin, M. Novitzky and J. J. Leonard (2016). Collision avoidance road test for COLREGS-constrained autonomous vehicles. OCEANS 2016 MTS/IEEE Monterey, IEEE.

Woods, D. and S. Dekker (2000). "Anticipating the effects of technological change: A new era of dynamics for human factors." Theoretical Issues in Ergonomics Science **1**(3): 272-282.

Wärtsilä (2019). Advanced intelligent Manoeuvring User Manual Version 2.90.5525.

Wärtsilä (2020). FOS AIM description – 2020. Internal Presentation

Zhang, R. and M. Furusho (2016). Constructing a Decision-Support System for Safe Ship-Navigation Using a Bayesian Network. Digital Human Modeling: Applications in Health, Safety, Ergonomics and Risk Management. V. G. Duffy. **9745:** 616-628.

Lighthouse brings together industry, society, academia and institutes in triple helix collaboration to strengthen Sweden's maritime competitiveness through research, development and innovation. As part of the work for a sustainable maritime sector, Lighthouse initiates and coordinates relevant research and innovation based on industry and society's needs.

Lighthouse – for a competitive, sustainable and safe maritime sector with a good working environment

LIGHTHOUSE PARTNERS

LIGHTHOUSE ASSOCIATE MEMBERS