

TRAVISIONS

TRA VISIONS 2022

Final Project Template – This template is to be used for submitting the final version of the project. The final deliverable should be submitted by the 31st of May 2022

Idea Number:

C2022-52

PILLAR / TRANSPORT MODE				
ROAD ☐	RAIL ☐	WATERBORNE ☒	AIRBORNE ☐	CROSS MODAL ☐
Project Related to:	☒ RA1: Smart Solutions and Society ☐ RA2: Green Mobility and Decarbonisation ☐ RA3: Innovative Infrastructure for Europe 2030 ☐ RA4: Policies and Economics for a Competitive Europe			

Project Title:

Trustworthy Autonomous Collision Avoidance Decision-Making System for Maritime Autonomous Surface Ships



This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement n° 101006593



1. Introduction

The recent maritime industry is paving the way towards the “Shipping 4.0” era [1], which is defined as a fully interconnected maritime ecosystem through the extensive use of cyber-physical and autonomous systems [2]. Autonomous ships, also known as the Maritime Autonomous Surface Ships (MASS), are envisioned to be the key enabling technologies for the fully autonomous maritime ecosystem, which according to the International Maritime Organisation are defined as ‘ships that to a varying degree can operate independently of human interaction’ [3]. Conventional ships are already the backbone of global economy that trade 90% of total goods’ volume [4], and the adoption of autonomous ones have the potential to achieve more sustainable and safer transportation [2]. However, a new emerging challenge that arises with the advancement of autonomous technologies is their trustworthiness, especially of those that execute mission-critical operations [5].

The aim of the proposed idea is to develop a trustworthy Autonomous Collision Avoidance Decision-Making System (ACA-DMS) for MASS. A reference system layout of the expected Autonomous Collision Avoidance System (ACAS) was conceptually developed and subsequently a methodology that combines the following tools/methods was employed to verify the feasibility of the proposed system: a) a digital twin of the system was developed to provide external risk metrics, such as proximity to obstacles using range detection sensors; b) Fault Tree Analysis (FTA) was conducted to acquire internal risk metrics, such as probability of failure of the system in various weather conditions; c) a reinforcement learning-based Deep Deterministic Policy Gradient (DDPG) agent was used to make collision avoidance decisions based on both external and internal risk metrics. A case cargo ship was employed considering specific operating areas to develop critical collision scenarios. The feasibility of the system was investigated by simulating the agent in trained and new operating areas prior unknown, and trustworthiness was evaluated based on both quantitative and qualitative metrics. The derived simulation results verified that the system was able to address collision scenarios of high-risk. In addition, it was verified that the consideration of both external and internal risk metrics led to the increase of the trustworthiness of the system. This idea is expected to have a great impact on the future maritime industry by providing a way forward towards safer, resilient, and more dependable trustworthy autonomous systems.

2. Selection of research area

Safe navigation at sea, especially the collision avoidance one, is deemed to be one of the most mission-critical autonomous operations for MASS [6]. The nature of the collision avoidance operation of a ship is a demanding endeavour of complex multi-tasking problems. It requires the control of an underactuated ship of large inertia on a dynamically changing and unstable liquid medium that is also exposed to unpredictable external factors, such as waves, currents, and various extreme weather conditions [7]. The adequate understanding of the manoeuvrability characteristics, traffic scenarios, navigational rules, and emergency scenarios require a great amount of cognitive ability and expert knowledge to generate sufficient collision avoidance decisions [7]. However, according to the European Maritime Safety Agency, recent accident statistics among European Union flag ships over 2014-2022 indicated that the accidents of navigational nature accounted for 43% of all occurrences with collision being the leading category, after contact and grounding [8].

The transition of the responsibilities from human operators to autonomous systems are expected to create additional challenges to the already demanding collision avoidance problem. Despite human error accounting for 94.7% of all collision accidents [9], the contribution of human operator in accident mitigation and prevention, despite not being reported, cannot be underestimated [10]. The decision-making of autonomous systems in unknown and emergency situations are hard to predict [11] and collision accidents due to system failure are expected to be the leading factor with the increase level of

autonomy [12]. Past experiences have already suffered the catastrophic consequences of collision accidents on global economy [13], environmental pollution [14], and life loss [9] and the extent of consequences in the autonomy spectrum is currently unknown.

The realisation of autonomous shipping, which is expected to achieve more sustainable and safer transportation whilst enhancing the resilience of the supply chain, is centric to the trustworthiness of the mission-critical autonomous systems, such as the ACA-DMS of MASS [5]. Trustworthiness is a basic quality of any system, especially of the mission-critical autonomous systems where the heterogeneity, interaction, and complexity of hardware, software, network, human factors, and physical environment is increased [15]. Trustworthiness is a multidimensional property that provides an accurate measurements of the holistic quality of a system [16], and it consists of various coherent attributes as presented in Figure 1. In other words, trustworthiness (T) can be defined as ‘a collective run-time quality of a system that performs a set of complex and interactive functions reliably and dependably’ [5], where $T = f(\text{safety, reliability, robustness, etc.})$. For each specific case study customisation is required to select the metrics that effectively characterise trustworthiness.

Definition 1. “Collision” is defined as the broader term any kind of contact that leads to the release of energy in the form of deformation, life loss, explosion, etc, including ship-ship striking, contact, or grounding.

Definition 2. “Collision avoidance” is defined as the actions that a ship takes to deviate from its original trajectory to avoid a potential collision (adapted from [17]).

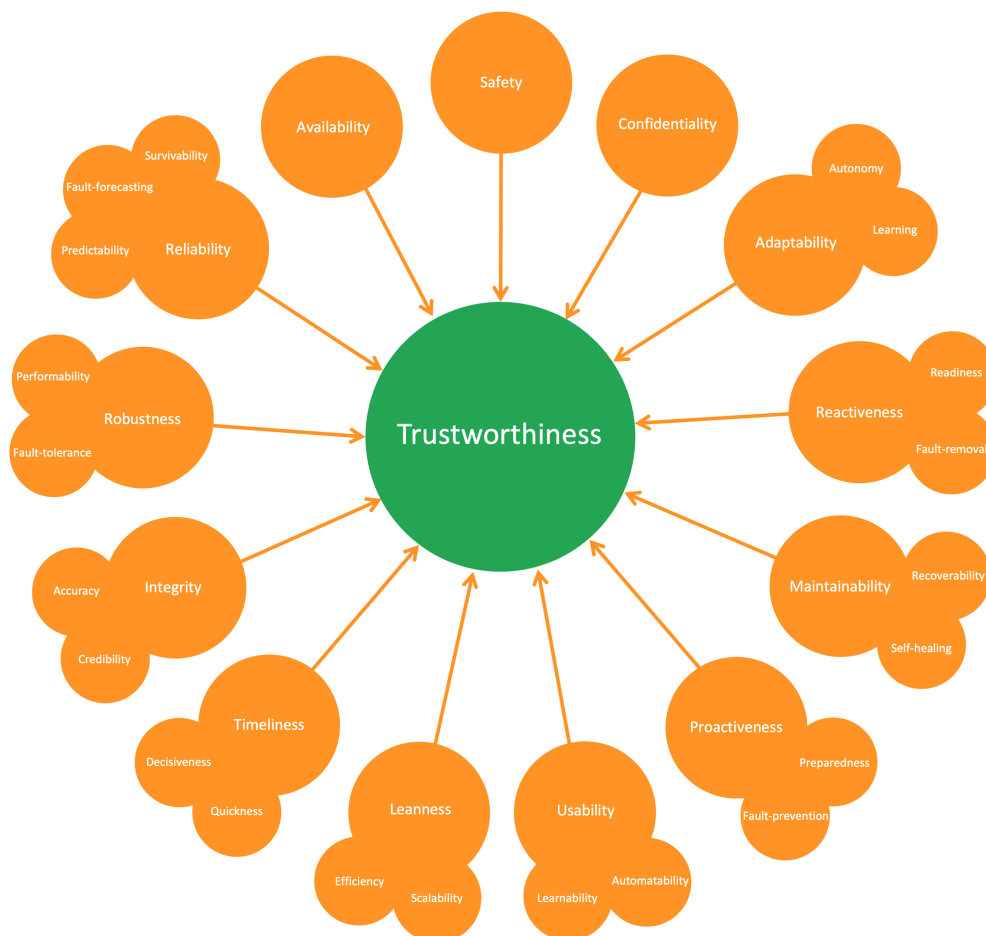


Figure 1 Trustworthiness of a system and its multi-disciplinary attributes with their equivalent attributes attached, according to Cho et al. [15].

3. Literature Survey

Trustworthiness of autonomous systems is an emerging concept that has not been rigorously explored in the literature [18]. According to Cho et al. [15] the key components to develop a trustworthy ACA-DMS are: 1) the investigated system with its functions, 2) the threats, such as errors and failures, 3) the attributes of trustworthiness, 4) the means to build a trustworthy system, and 5) the relationship between the trustworthiness attributes and assessment. For each key component a critical literature review has been conducted and the summary of the main findings is presented in Table 1.

In order to investigate the safety of MASS, recent risk analysis has focused in analysing the risk associated with the design and operation of MASS. Rødseth et al. [6] and Thieme et al. [19] conducted Preliminary Hazard Analysis (PHA) to realise a risk-informed design of an unmanned ships and small autonomous ferry, respectively. Banda et al. [20] and Wrobel et al. [21] did Systems-Theoretic Process Analysis (STPA) to identify the control structure of autonomous ships. Wróbel et al. [22] used What-If analysis, Human Factor Analysis, and Classification System for Marine Accident, and Wu et al. [10] used Hybrid Causal Logic (HCL) to analyse the risk associated with the operation of unmanned ships compared to conventional maritime accidents. However, previous works were limited to qualitatively analysing the risks associated with the generic system and operations, and quantitative analysis of mission-critical autonomous systems, such as the ACAS, was not found.

In order to investigate the feasibility of MASS, a new body of research has been emerged that use state-of-the-art machine learning techniques, such as Deep Reinforcement Learning (DRL) algorithms, in order to develop autonomous decision-making systems, such as ACA-DMS. Wang et al. [23] and Zhang et al. [24] used Deep Q-Learning (DQL) algorithm to generate collision avoidance decisions in static and dynamic obstacles. Cheng et al. [25] and Martinsen et al. [26] used a DRL algorithm, such as DDPG, to control the collision avoidance in static environment under unknown environmental disturbances and sea currents, respectively. Zhao et al. [27] used Proximal Policy Optimization (PPO) algorithm to realise path planning and collision avoidance among other moving ships, while performing International Regulations for Preventing Collisions at Sea (COLREG) rules compliance and control under unknown environmental disturbances such as wave, current, and wind. However, the previously developed decision-making systems generated decisions solely based on external risk metrics and decision-making based on both internal and external risk metrics were not found.

The literature review outcomes demonstrate that the major gap identified is the following: ACA-DMS has made decisions solely based on external risk metrics and internal risk metrics have not been considered, which consequently degrades the quality of many attributes of the trustworthiness of the system. Thus, the aim of the proposed concept is to develop a trustworthy ACA-DMS that generates collision avoidance decisions based on both critical external and internal risk metrics in order to enhance the quality of its trustworthiness.

4. Proposed Solution-Description of Idea

The methodology flowchart of the proposed concept idea consists of 5 main phases, as presented in Figure 2. Phase 1 includes the design of a preliminary reference system layout of an ACAS that is capable of performing the autonomous collision avoidance operations. The proposed system provides the reference for the development of a digital twin in order to investigate the trustworthiness of the system in a virtual environment by utilising Matlab/Simulink. Further information of the details and rationale of the reference system design is presented in Subsection 6.1. Reference system layout.

[illegible]

In Phase 3, risk analysis of the digital twin is conducted using FTA. More specifically, various weather variations based on good & bad weather and day & night conditions are considered and, consequently, 4 different system layouts for each weather condition are developed. A fault tree that models the failure propagation leading to the failure of the whole system is developed for each system layout. Numerical values, such as Mean Time Between Failure (MTBF), for each component are acquired and used as input data to the analysis. The analysis of each fault tree provides quantitative internal risk metrics, such as component's importance measures and probability of failure of the whole system in each weather condition as a function of time. Additional information and supporting details are presented in Subsection 6.2. Fault tree analysis.

Phase 4 employs an artificial intelligence agent which is used to develop the ACA-DMS that make collision avoidance decisions based on both external and internal risk metrics generated in the previous Phases 2-3. For this reason, a DRL based DDPG algorithm is employed by setting up the appropriate hyperparameters and reward functions. The developed agent is then trained to make collision avoidance decisions by giving the appropriate rudder angle and propeller revolution commands. The main design characteristics of the agent are described in Subsection 6.3. Intelligent decision-making agent.

Finally, the trustworthiness of the ACA-DMS is investigated and verified in Phase 5. The fully trained agent is simulated in various critical scenarios, including previously trained and prior unknown ones. The feasibility of the developed system to control the ship and successfully performance collision avoidance are investigated during the simulations. Ultimately, the effect of considering both critical external and internal risk metrics on the quality of the trustworthiness of ACA-DMS is evaluated based on qualitative and quantitative metrics.

5. Address potential technology gaps

5.1. Case scenarios

To investigate the feasibility of the proposed idea, two operating areas are selected considering an autonomous cargo ship as presented in Figure 3. The first one considers a highly restricted sea area of 2 x 4 km with static obstacles that represent shore areas, small islands, and anchored ships. The investigated ship is initiated at the South side with the aim of exiting through the North pathways by avoiding the obstacles. The second one considers an open sea area of 3 x 3 km and another moving ship. The investigated ship is initiated at the South side with the aim of exiting North by successfully executing the approaching heading scenario according to the CORELG rules. In both areas good weather & day condition is considered.

Each operating areas accommodates its own challenges. In the first area, the challenge is not only to successfully navigate through the narrow path by not colliding but also to exit from the safest pathway at the North side by choosing between the wider and narrower pathways. In the second one, the challenge is to successfully comply to the heading COLREG rule while the other ship is violating it by not changing its heading route.

Another challenge common to both areas is the fact that the digital twin does not have information of the whole scene but has situation awareness within the defined range of 1500 m. The zone outside of the detectable area is considered to be the "blind zone" where no information can be obtained. Considering the large inertia of the ship and relatively short situation awareness range and time to react, both areas are of high-risk, where the decision-making capabilities can be tested to its limits.

In total of 4 different case scenarios are developed by investigating cases where only external risk metrics and both external and internal risk metrics are considered for each

operating area. In the latter case, the probability of failure of the digital twin is defined as the critical internal risk metric that represents the health condition of the whole system as a function of time. In the former case, a static ship domain is considered where no interference with other obstacles is desired, whereas a dynamic ship domain that increases its size proportionate to the increase of the probability of failure is considered in the latter case.

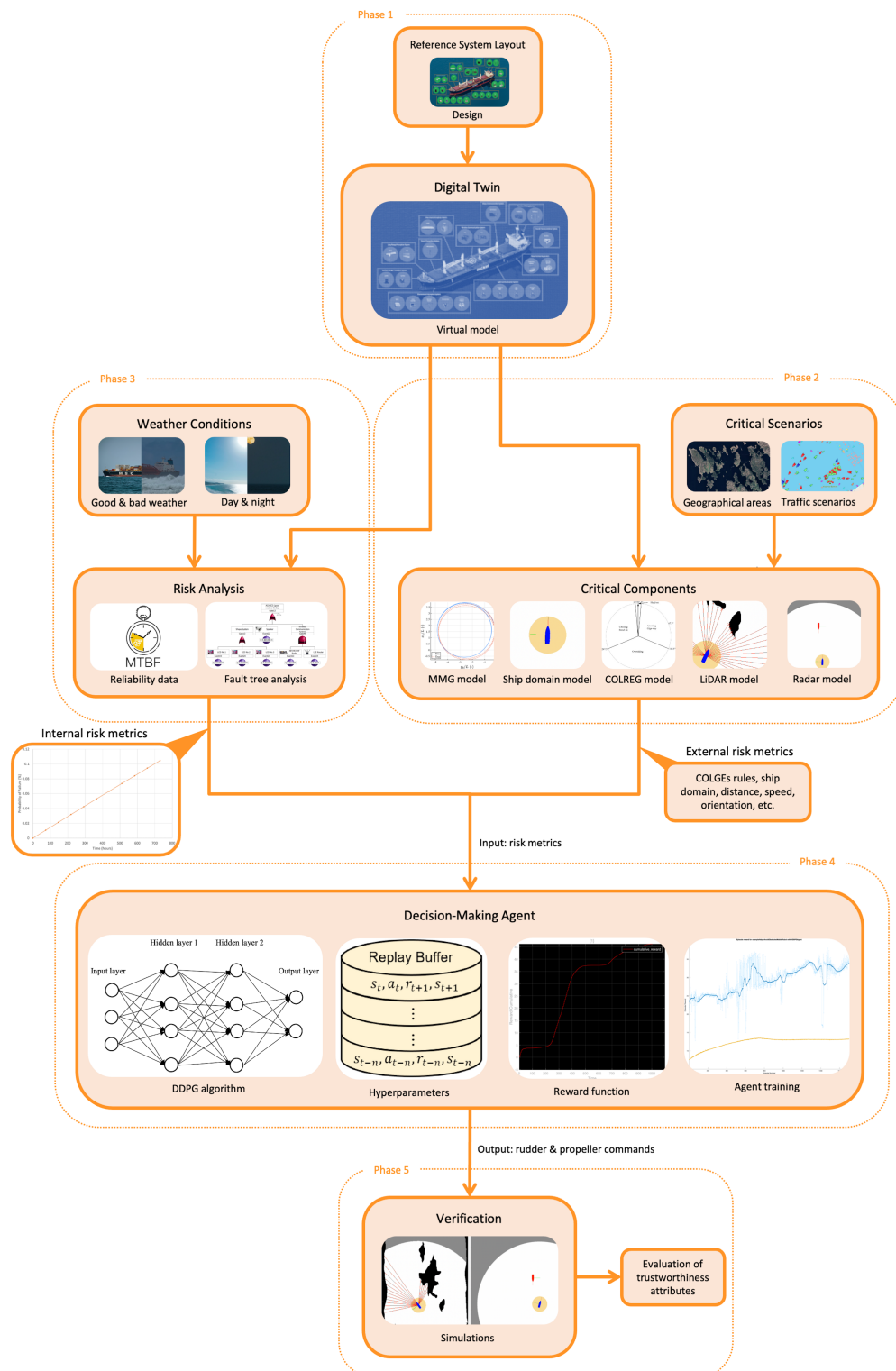


Figure 2 Proposed methodology flowchart.

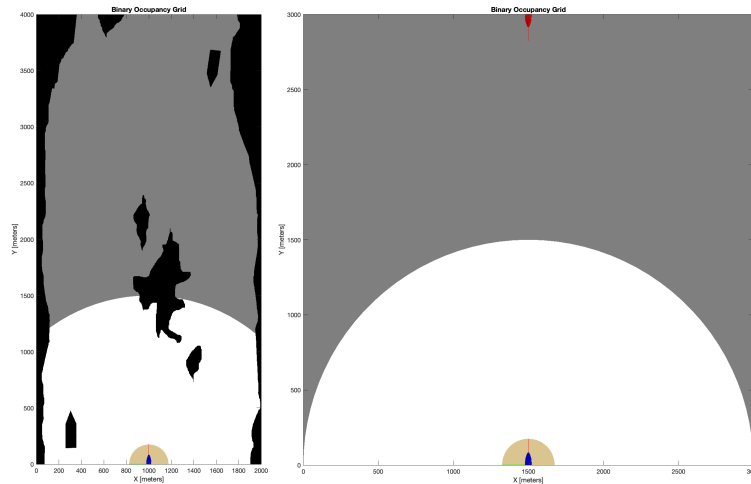


Figure 3 The two operating areas selected for training: collision avoidance in restricted sea area with static obstacles (left) and in open sea area with dynamic obstacles (right).

5.2. Results & Discussion

The derived simulation results verified the feasibility of the develop ACA-DMS, as presented in Figure 4-Figure 5. In all scenarios the agent was able to avoid collisions, not violate the ship domain, and not conduct unnecessary extra manoeuvres, such as circle manoeuvres, by giving the appropriate rudder angle and propeller revolution commands, as presented in Figure 6. More particularly, in the restricted area the agent was able choose the safest exit by manoeuvring through the wider pathway. In the open environment the agent was able to avoid collision in a heading situation by complying to the COLREG rules while the other ship was violating it. Also, the generalisation capabilities of the agent were investigated by simulating it in new operating areas prior unknown, as presented in Figure 7. Even these new areas, the agent was able to generalise the knowledge acquired from previous trainings and successfully perform collision avoidance.

In addition, the attributes of trustworthiness were evaluated based on quantitative and qualitative metrics as presented in Table 2. From the derived simulation results it can be derived that the developed ACA-DMS was able achieve five of the main attributes of trustworthiness, including: a) integrity due to its accuracy in performing the collision avoidance actions; b) adaptability due to its autonomy to make decisions without human intervention; c) usability due to its learnability to control the ship by giving the appropriate rudder angle and propeller revolution commands; d) timeliness due to its decisiveness to generate collision avoidance decisions based on real-time observations; and, e) leanness due to its efficiency to conduct only the required manoeuvres. Also, it can be argued that safety was also achieved at a lower level due to its ability to avoid collisions, not violate ship domain, and comply to COLREG rules.

Finally, it can be noted that the ACA-DMS was able to further enhance the quality of trustworthiness by enhancing previous attributes and achieving new ones, when both external and internal risk metrics were considered. For instance, safety was further enhanced as the system became less risk-appetite by proceeding with smaller sailing speeds and conducting less manoeuvres, as presented in Figure 5-Figure 6, which consequently affected leanness as well. A new attribute that was achieved was reliability due to its ability to fault-forecast the system's health condition. Also, robustness and proactiveness were enhanced due to the fault-tolerant decisions amidst the system's degradation.

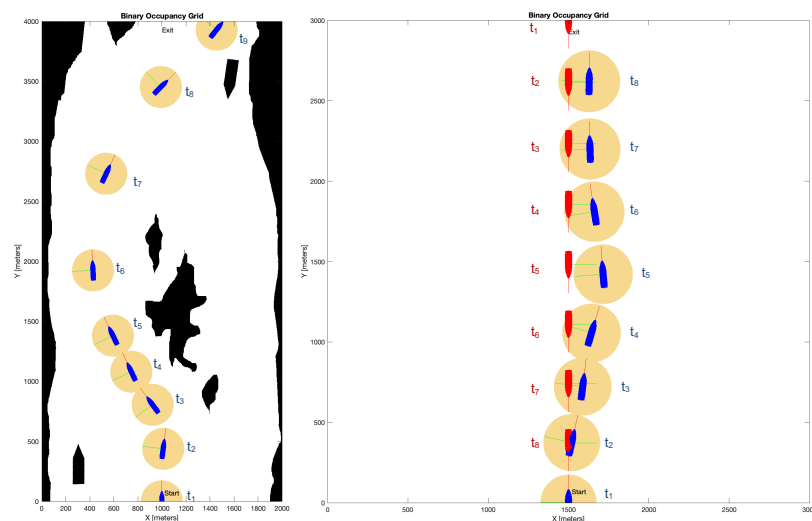


Figure 4 Simulation results of the trained agents that consider only external risk metrics in both operating areas (t : time step).

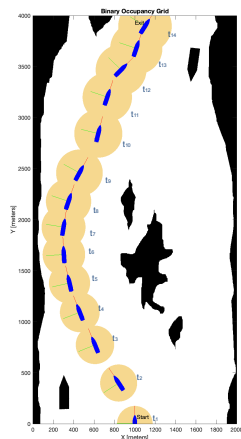


Figure 5 Simulation results of the trained agent that consider both external and internal risk metrics in the restricted operating area (t : time step).

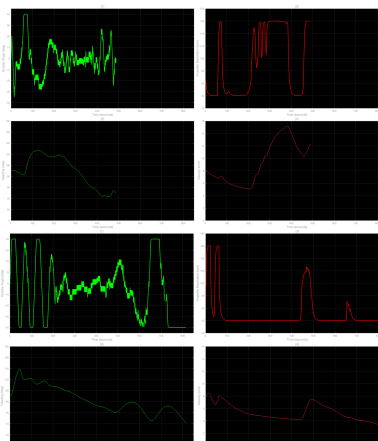


Figure 6 Rudder angle and propeller revolution commands of the agents, including heading angle, and velocity in the restricted operating area considering only external risk metrics (top) and considering both external and internal risk metrics (bottom).

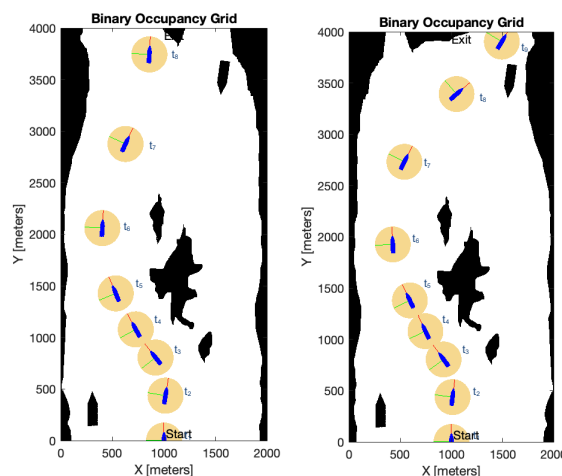


Figure 7 Generalisation capabilities of the trained agent: the agent is able to generalise the knowledge acquired from previous trainings and perform collision avoidance even in new operating areas prior unknown (t : time step).

Table 2 Qualitative evaluation of the main attributes of trustworthiness for each scenario. Heatmap: ● is used to represents “Not achieved”, ● is used to represent “Achieved”, and ● is used to represent “Enhanced”.

Main attributes of trustworthiness	Internal risk metrics not considered		Internal risk metrics considered	
	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Reliability	●	●	●	●
Availability	●	●	●	●
Safety	●	●	●	●
Confidentiality	●	●	●	●
Integrity	●	●	●	●
Robustness	●	●	●	●
Maintainability	●	●	●	●
Adaptability	●	●	●	●
Usability	●	●	●	●
Timeliness	●	●	●	●
Leanness	●	●	●	●
Reactivity	●	●	●	●
Proactiveness	●	●	●	●

To support the aforementioned observations, various quantitative metrics were calculated. For instance, sailing velocity and proximity to obstacles are of the most basic quantitative metrics that need to be evaluated as lower velocity and greater proximity from obstacles consequently leads to increased safety from collision and enhanced robustness and proactiveness against system degradation and occurrences of critical failures. Histogram charts of the velocity and proximity to obstacles distributions pertaining to the cases of restricted sea area were generated in order to evaluate the frequency of data, as presented in Figure 8. Also, median values of each were calculated as quantitative metrics for trustworthiness: $M_{vel}^{ext} = 7.91 \text{ m/s}$, $M_{vel}^{int+ext} = 4.93 \text{ m/s}$, $M_{prox}^{ext} = 1089.21 \text{ m}$, and $M_{prox}^{int+ext} = 1179.61 \text{ m}$. Quantitative metrics also support the idea that the consideration of both external and internal risk metrics lead to the increase of trustworthiness of system.

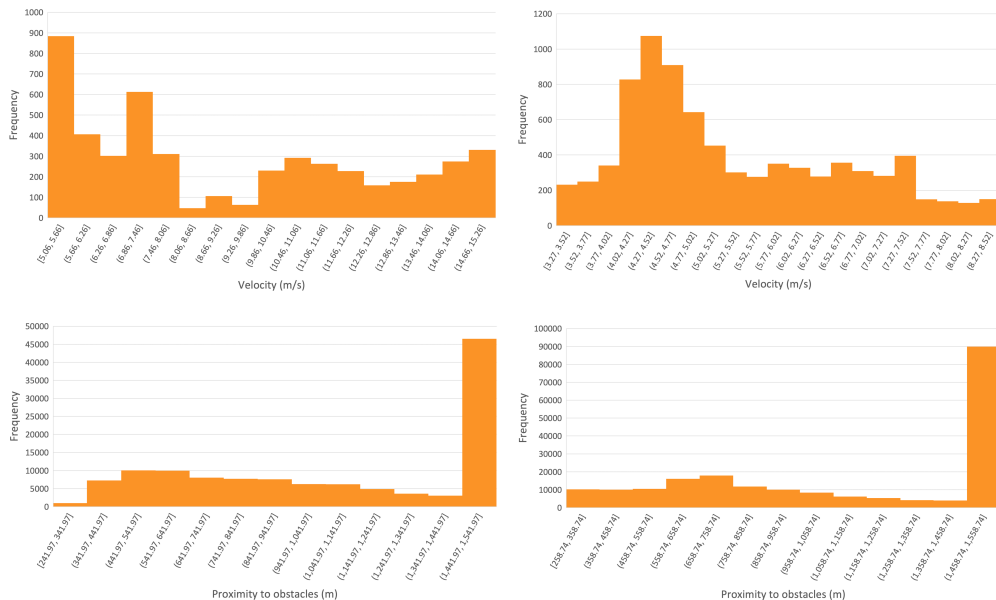


Figure 8 Histogram charts of the velocity and proximity to obstacles regarding the case of considering only external risk metrics (left side) and considering both internal and external risk metrics (right side).

5.3. Conclusions & Future research

The proposed idea develops a trustworthy ACA-DMS for MASS. The main findings are summarised as follows:

- The developed system, ACA-DMS, is feasible for the control and decision-making in complex and critical autonomous collision avoidance operations.
- The ACA-DMS have great generalisation capabilities, suitable for real-time operation and in new environments prior unknown.
- The consideration of external risk metrics to the decision-making provides a certain level of trustworthiness of the ACA-DMS by achieving some of its main attributes.
- The consideration of both external and internal risk metrics is pivotal to further enhance the trustworthiness of ACA-DMS, which is verified based on both qualitative and quantitative metrics.

This idea is expected to have a great impact with potentially direct implications to the current and future maritime transportation, as it is applicable to both conventional and next generation autonomous ships and expandable for more high-risk collision scenarios. The utilisation of artificial intelligence and machine learning methods that are well-risk informed will enable the execution of the complex collision avoidance without human interference, with the potential of performing at the same or even higher cognitive level than humans. This will potentially lead to the implementation of safer MASS and towards the realisation of net zero safety maritime transportation through net zero accidents. Suffice it to say, a fully autonomous shipping of such trustworthy MASS will also aid towards the realisation of net zero emission too.

It is pivotal to mention that the proposed idea has its limitations with potential for further developments. Some of the mentionable limitations are the limited scenarios examined, the difficulty of identifying the most critical and unpredictable scenario, such as unknown unknowns, and verification of the developed system. Other limitations include the limitation of the considered critical external and internal risk metrics to tackle and enhance all attributes of trustworthiness, the ambiguity of the qualitative metrics, and the lack of a quantitative metric to evaluate trustworthiness of a system holistically. Future developments include the identification of more critical scenarios, consideration of more critical internal risk metrics, such as uncertainty of input data from the range detection sensors and more realistic distribution of MTBF for each component to consider the extreme failure rates of the most critical components. Other future works include the development of a wholistic quantitative metric to better evaluate trustworthiness and development of a more complex digital twin that considers more complex interactions and threats, such as cyber-security, in order to investigate more attributes of trustworthiness.

6. Detailed design description

6.1. Reference system layout

The expected ACAS is required to be equipped with all the necessary onboard components and systems that allow the execution of autonomous collision avoidance operations, which are: a) situation awareness (using radar, camera, AIS, LiDAR, GPS, etc.), decision-making (main computer hardware and software), communication (using speaker, navigational lights, LTE, etc.), and execution (using power generation system, steering system, etc.). A generic reference system layout is presented in Figure 9.



A general fault tree diagram that models the failure propagation from the basic components level to the failure of the ACAS in one of the investigated weather conditions as the top event is presented in Figure 10. “OR” gates were used to model components with critical causal relationship, “AND” to model complimentary functionality, and “VOTE” gate for redundancy. MTBF of the components are presented in Table 3. Results, such as the probability of failure of the top event, can be derived as presented in Figure 11.



Table 3 MTBF data of the ACAS components.

Subsystem	Component (Product Name)	MTBF (hours)
Situation Awareness System	INS (MK31) [48]	16,000
	GNSS (SEANAV 300 Series) [49]	45,000
	Day/night & Thermal Camera (CLRT/HD-400) [50]	20,200
	LiDAR (M1 Edge) [51]	60,000
	Radar (REKA) [52]	87,600
	AIS (AIS 300S) [53]	100,000
	Echo Sounder (NESDF) [54]	20,000
	Visibility & Present Weather Sensor (VPF-730) [55]	56,700
	Anemometre (Windsonic 75) [56]	131,400
	VDL & ADCP (Syrinx) [57]	225,000
	ECDIS (MPC-122X-K-ECDIS) [58]	39,675
Intention Communication System	Microphone (UA-1404) [59]	40,000
	Masthead light [60]	30,000
	Starboard-Side light [60]	30,000
	Port-Side light [60]	30,000
	All-Round light [60]	30,000
	Speaker (VML-1520) [61]	1,500,000
	LCD (DLF/DLH1968-U) [62]	70,000
	MF/HF/VHF Radio [63]	50,000
Decision Making System	LTE Router (RUT950) [64]	270,000
	Computer Hardware (LEVIATHAN 17i FANLESS) [65]	100,000
Action Execution System	Software [66]	40,000
	Power & Propulsion System [67]	42,048

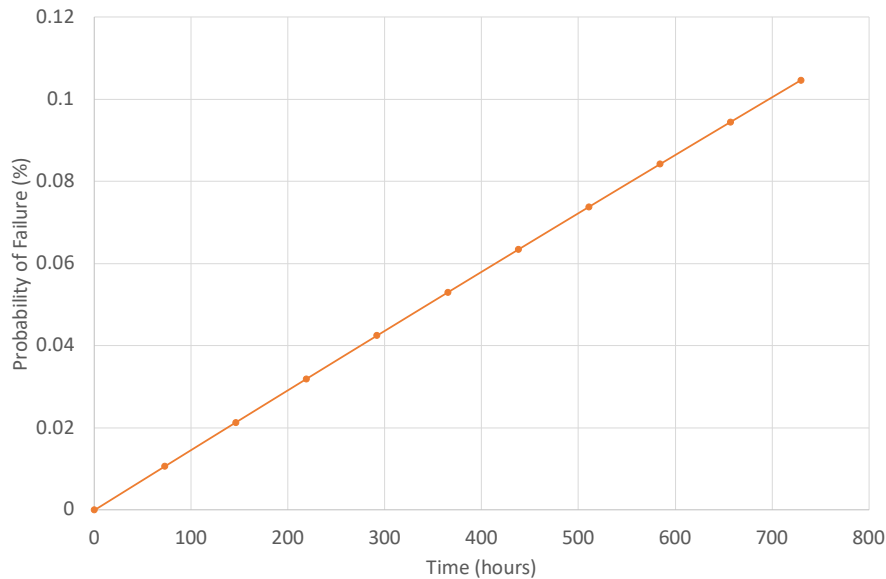


Figure 11 Probability of failure of the ACAS in good weather & day.

6.3. Intelligent decision-making agent

The DDPG algorithm consists of two deep neural networks, known as the actor network and the critic network. The actor network approximates the action from the current state, while the critic network updates the parameters of the actor network by generating a Q-value, which is an estimate of the discounted long-term reward at the start of each episode, given the initial observation of the environment [25].

The actor network is customised to have one input layer from the environment observations, two hidden layers with 400 units each, and an output layer that produces the action commands. The critic network has one input layer from the environment observations and action commands, two hidden layers in between with 300 units each, and an output layer that generates the Q-value. The input consists of the ship states, including the real-time x-position, y-position, heading angle, and the proximity measurements from the distance detection sensors, such as lidar sensors. The output

consists of the action commands regarding the rudder angle and the propeller's revolution.

The hyperparameters used are 10^{-3} learning rate for the critic network, 10^{-4} learning rate for the actor network, and 256 batch size. The developed reward functions are the following:

- Reward for heading North:

$$r_{heading} = \frac{1}{(Heading - 90^\circ)^2 + 1}$$

- Reward for keeping obstacles outside the ship domain:

$$r_{ship\ domain} = \begin{cases} -1000, & proximity < ship\ domain \\ 10^{-5} * proximity, & proximity \geq ship\ domain \end{cases}$$

- Reward for complying to the COLREG rules:

$$r_{Heading\ COLREG} = \begin{cases} -1000, & relative\ angle \leq 0 \\ 10^{-3} * relative\ angle, & relative\ angle > 0 \end{cases}$$

7. Source of additional information used & Appendices

7.1. Abbreviations

Table 4 Abbreviations used.

Abbreviation	Full form
AC	Actor-Critic
ACA-DMS	Autonomous Collision Avoidance Decision-Making System
ACAS	Autonomous Collision Avoidance System
ADCP	Acoustic Doppler Current Profiler
AIS	Automatic Identification System
BNN	Bayesian Network
CFD	Computational Fluid Dynamics
COLREG	International Regulations for Preventing Collisions at Sea
DCPA	Distance at Closet Point of Approach
DDPG	Deep Deterministic Policy Gradient
DRL	Deep Reinforcement Learning
DQL	Deep Q-Learning
DQN	Deep Q-Network
ECDIS	Electronic Chart Display and Information System (ECDIS)
EMSA	European Maritime Safety Agency
FTA	Fault Tree Analysis
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HCL	Hybrid Causal Logic
IMO	International Maritime Organisation
INS	Inertial Navigation System
LCD	Liquid-Cristal Display
LIDAR	Light Detection and Ranging
LTE	Long Term Evolution
MASS	Maritime Autonomous Surface Ships
MMG	Manoeuvring Modelling Group
MTBF	Mean Time Between Failure
PG	Policy Gradient
PHA	Preliminary Hazard Analysis
PPO	Proximal Policy Optimization
Q	Q-Learning
Radar	Radio Detection and Ranging
SAC	Soft Actor-Critic
STPA	Systems Theoretic Process Analysis
SWIFT	Structured What-If Technique
TCPA	Time to Closet Point of Approach
TD3	Twin-Delayed Deep Deterministic Policy Gradient
TRPO	Trust Region Policy Optimisation
VDL	Velocity Doppler Log
VHF	Very High Frequency

8. References

- [1] Ø. J. Rødseth, L. P. Perera, and B. Mo, "Big data in shipping-Challenges and opportunities," 2016.
- [2] G. Aiello, A. Giallanza, and G. Mascarella, "Towards Shipping 4.0. A preliminary gap analysis," *Procedia Manufacturing*, vol. 42, pp. 24-29, 2020.
- [3] IMO. "IMO takes first steps to address autonomous ships."
<https://www.imo.org/en/MediaCentre/PressBriefings/Pages/08-MSC-99-MASS-scoping.aspx#:~:text=For%20the%20purpose%20of%20the,operate%20independentl y%20of%20human%20interaction.&text=Remotely%20controlled%20ship%20without %20seafarers.and%20operated%20from%20another%20location.> (accessed.
- [4] R. Asariotis *et al.*, "Review of Maritime Transport, 2011," 9211128412, 2011.
- [5] Y. Wang *et al.*, "A tripartite theory of trustworthiness for autonomous systems," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2020: IEEE, pp. 3375-3380.
- [6] Ø. Rødseth and A. Tjora, "A risk based approach to the design of unmanned ship control systems," in *Proceeding Conf Marit Technol*, 2014, pp. 153-62.
- [7] T. Statheros, G. Howells, and K. M. Maier, "Autonomous ship collision avoidance navigation concepts, technologies and techniques," *The Journal of Navigation*, vol. 61, no. 1, pp. 129-142, 2008.
- [8] EMSA, "Annual Overview of Marine Casualties and Incidents 2021," 2021. [Online]. Available: <http://www.emsa.europa.eu/newsroom/latest-news/item/4266-annual-overview-of-marine-casualties-and-incidents-2020.html>
- [9] H. Ugurlu and I. Cicek, "Analysis and assessment of ship collision accidents using Fault Tree and Multiple Correspondence Analysis," *Ocean Engineering*, vol. 245, p. 110514, 2022.
- [10] Q. Wu, T. Wang, M. A. Diaconeasa, A. Mosleh, and Y. Wang, "A Comparative Assessment of Collision Risk of Manned and Unmanned Vessels," *Journal of Marine Science and Engineering*, vol. 8, no. 11, p. 852, 2020.
- [11] M. Rausand and S. Haugen, *Risk assessment: theory, methods, and applications*, Second ed. John Wiley & Sons, 2020.
- [12] B. Veritas, "NI641 guidelines for autonomous shipping," ed, 2019.
- [13] J. M.-y. Lee and E. Y.-c. Wong, "Suez Canal blockage: an analysis of legal impact, risks and liabilities to the global supply chain," in *MATEC Web of Conferences*, 2021, vol. 339: EDP Sciences.
- [14] J. Chen *et al.*, "Marine oil spill pollution causes and governance: A case study of Sanchi tanker collision and explosion," *Journal of Cleaner Production*, vol. 273, p. 122978, 2020.
- [15] J.-H. Cho, P. M. Hurley, and S. Xu, "Metrics and measurement of trustworthy systems," in *MILCOM 2016-2016 IEEE Military Communications Conference*, 2016: IEEE, pp. 1237-1242.
- [16] J.-H. Cho, S. Xu, P. M. Hurley, M. Mackay, T. Benjamin, and M. Beaumont, "Stram: Measuring the trustworthiness of computer-based systems," *ACM Computing Surveys (CSUR)*, vol. 51, no. 6, pp. 1-47, 2019.
- [17] Y. Huang, L. Chen, P. Chen, R. R. Negenborn, and P. van Gelder, "Ship collision avoidance methods: State-of-the-art," *Safety science*, vol. 121, pp. 451-473, 2020.
- [18] H. A. Abbass, G. Leu, and K. Merrick, "A review of theoretical and practical challenges of trusted autonomy in big data," *IEEE Access*, vol. 4, pp. 2808-2830, 2016.
- [19] C. A. Thieme, C. Guo, I. B. Utne, and S. Haugen, "Preliminary hazard analysis of a small harbor passenger ferry—results, challenges and further work," in *Journal of Physics: Conference Series*, 2019, vol. 1357, no. 1: IOP Publishing, p. 012024.
- [20] O. A. V. Banda, S. Kannos, F. Goerlandt, P. H. van Gelder, M. Bergström, and P. Kujala, "A systemic hazard analysis and management process for the concept design phase of an autonomous vessel," *Reliability Engineering & System Safety*, vol. 191, p. 106584, 2019.
- [21] K. Wróbel, P. Krata, and J. Montewka, "Preliminary results of a system-theoretic assessment of maritime autonomous surface ships' safety," *TransNav, International Journal on Marine Navigation and Safety od Sea Transportation*, vol. 13, no. 4, 2019.

- [22] K. Wróbel, J. Montewka, and P. Kujala, "Towards the assessment of potential impact of unmanned vessels on maritime transportation safety," *Reliability Engineering & System Safety*, vol. 165, pp. 155-169, 2017.
- [23] C. Wang, X. Zhang, L. Cong, J. Li, and J. Zhang, "Research on intelligent collision avoidance decision-making of unmanned ship in unknown environments," *Evolving Systems*, vol. 10, no. 4, pp. 649-658, 2019.
- [24] X. Zhang, C. Wang, Y. Liu, and X. Chen, "Decision-making for the autonomous navigation of maritime autonomous surface ships based on scene division and deep reinforcement learning," *Sensors*, vol. 19, no. 18, p. 4055, 2019.
- [25] Y. Cheng and W. Zhang, "Concise deep reinforcement learning obstacle avoidance for underactuated unmanned marine vessels," *Neurocomputing*, vol. 272, pp. 63-73, 2018.
- [26] A. B. Martinsen and A. M. Lekkas, "Curved path following with deep reinforcement learning: Results from three vessel models," in *OCEANS 2018 MTS/IEEE Charleston*, 2018: IEEE, pp. 1-8.
- [27] L. Zhao, M.-I. Roh, and S.-J. Lee, "Control method for path following and collision avoidance of autonomous ship based on deep reinforcement learning," *Journal of Marine Science and Technology*, vol. 27, no. 4, p. 1, 2019.
- [28] K. Nomoto, "Analysis of Kempfs standard maneuver test and proposed steering quality indices," in *Osaka University, Japan, Presented at: Proceedings of the 1st Symposium on Ship Manoeuvrability, David Taylor Model Basin, Washington DC, USA, Report 1461*, pp. 275-305, 1960.
- [29] M. Zhu, A. Hahn, Y.-Q. Wen, and W.-Q. Sun, "Optimized support vector regression algorithm-based modeling of ship dynamics," *Applied Ocean Research*, vol. 90, p. 101842, 2019.
- [30] M. A. Abkowitz, "Lectures on ship hydrodynamics--Steering and manoeuvrability," 1964.
- [31] E. Kobayashi, H. Kagemoto, and Y. Furukawa, "Research on ship manoeuvrability and its application to ship design. Chapter 2: mathematical models of manoeuvring motions," in *The 12th Marine Dynamic Symposium*, 1995, pp. 23-90.
- [32] R. G. Wright, "Intelligent autonomous ship navigation using multi-sensor modalities," *TransNav: International Journal on Marine Navigation and Safety of Sea Transportation*, vol. 13, no. 3, 2019.
- [33] A. Felski and K. Zwolak, "The Ocean-Going Autonomous Ship—Challenges and Threats," *Journal of Marine Science and Engineering*, vol. 8, no. 1, p. 41, 2020.
- [34] M. Höyhty, J. Huusko, M. Kiviranta, K. Solberg, and J. Rokka, "Connectivity for autonomous ships: Architecture, use cases, and research challenges," in *2017 International Conference on Information and Communication Technology Convergence (ICTC)*, 2017: IEEE, pp. 345-350.
- [35] C.-H. Chang, C. Kontovas, Q. Yu, and Z. Yang, "Risk assessment of the operations of maritime autonomous surface ships," *Reliability Engineering & System Safety*, vol. 207, p. 107324, 2021.
- [36] K. Wróbel, J. Montewka, and P. Kujala, "System-theoretic approach to safety of remotely-controlled merchant vessel," *Ocean Engineering*, vol. 152, pp. 334-345, 2018.
- [37] H.-C. Burmeister, W. C. Bruhn, Ø. J. Rødseth, and T. Porathe, "Can unmanned ships improve navigational safety?," in *Proceedings of the Transport Research Arena, TRA 2014, 14-17 April 2014, Paris*, 2014.
- [38] H.-C. Burmeister, W. Bruhn, Ø. J. Rødseth, and T. Porathe, "Autonomous unmanned merchant vessel and its contribution towards the e-Navigation implementation: The MUNIN perspective," *International Journal of e-Navigation and Maritime Economy*, vol. 1, pp. 1-13, 2014.
- [39] Y. Man, R. Weber, J. Cimbritz, M. Lundh, and S. N. MacKinnon, "Human factor issues during remote ship monitoring tasks: An ecological lesson for system design in a distributed context," *International Journal of Industrial Ergonomics*, vol. 68, pp. 231-244, 2018.
- [40] K. Wróbel, J. Montewka, and P. Kujala, "Towards the development of a system-theoretic model for safety assessment of autonomous merchant vessels," *Reliability Engineering & System Safety*, vol. 178, pp. 209-224, 2018.

- [41] C. A. Thieme and I. B. Utne, "A risk model for autonomous marine systems and operation focusing on human–autonomy collaboration," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 231, no. 4, pp. 446-464, 2017.
- [42] Matlab. "Reinforcement learning agents." <https://www.mathworks.com/help/reinforcement-learning/ug/create-agents-for-reinforcement-learning.html> (accessed).
- [43] J. Woo and N. Kim, "Collision avoidance for an unmanned surface vehicle using deep reinforcement learning," *Ocean Engineering*, vol. 199, p. 107001, 2020.
- [44] D.-H. Chun, M.-I. Roh, H.-W. Lee, J. Ha, and D. Yu, "Deep reinforcement learning-based collision avoidance for an autonomous ship," *Ocean Engineering*, vol. 234, p. 109216, 2021.
- [45] Y. He, Y. Jin, L. Huang, Y. Xiong, P. Chen, and J. Mou, "Quantitative analysis of COLREG rules and seamanship for autonomous collision avoidance at open sea," *Ocean Engineering*, vol. 140, pp. 281-291, 2017.
- [46] Y. Huang, L. Chen, R. R. Negenborn, and P. Van Gelder, "Collision avoidance systems for maritime autonomous surface ships considering uncertainty in ship dynamics," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 14614-14619, 2020.
- [47] X. Wu, K. Liu, J. Zhang, Z. Yuan, J. Liu, and Q. Yu, "An Optimized Collision Avoidance Decision-Making System for Autonomous Ships under Human-Machine Cooperation Situations," *Journal of Advanced Transportation*, vol. 2021, 2021.
- [48] Teledynemarine. "MK31 INS." <http://www.teledynemarine.com/Lists/Downloads/TSS/TSS%20MK31%20INS%20Product%20leaflet.pdf> (accessed).
- [49] Kongsberg. "Seanav 300 Series." https://www.kongsberg.com/contentassets/0d2751ec14f74850bf4106ace4472d5f/datasheet_seanav300.pdf (accessed).
- [50] DatCon, "Long and Medium Range - Modular Multi Sensor System," 2022. [Online]. Available: <https://www.dat-con-defence.com/wp-content/uploads/2020/07/Modular-multi-sensor-system-1.pdf>.
- [51] Quanergy. "M1 Edge - Smart 2D LiDAR Solution." <https://quanergy.com/products/m1-edge/> (accessed).
- [52] Micran. "Radio navigation system REKA." <http://micran.com/productions/radar-systems/radio-navigation-system-reka/> (accessed).
- [53] Kongsberg. "Automatic Identification System - Special Operations: AIS 300S." https://www.kongsberg.com/contentassets/abc742e411b04f1ab77096bc04194195/datasheet_ais300s.pdf (accessed).
- [54] NautelSonar. "Naval echo sounder - NESDF." <https://www.nautelsonar.com/wp-content/uploads/2019/01/Nautel-Sonar-NES-DF.pdf> (accessed).
- [55] MicroStep-MIS. "Visibility and present weather sensor: VPF-730 Series " https://www.microstep-mis.com/drupal/web/sites/default/files/datasheets/VPF-730_product_sheet.pdf (accessed).
- [56] MicroStep-MIS. "Windsonic 75 Anemometer." https://www.microstep-mis.com/web/products?category=ultrasonic&product=windsonic_75_anemometer (accessed).
- [57] Sonardyne. "Doppler Velocity Log with integrated ADCP - Syrinx DVL." <https://www.sonardyne.com/products/syrinx-doppler-velocity-log/> (accessed).
- [58] DigitX. "MPC-122X-K-ECDIS." https://www.digitx.it/codice/mpc_122x_k_ecdis/brand-moxa/id_prod-92338/ (accessed).
- [59] BrüelKjær. "Outdoor microphone kit - UA-1404." <https://www.bksv.com/doc/bp1696.pdf> (accessed).
- [60] Oxleygroup. "Marine Navigation Lighting and Control Systems." <https://www.oxleygroup.com/images/newproductspdf/189.pdf> (accessed).
- [61] Zenitel. "MTBF Figures." https://wiki.zenitel.com/wiki/MTBF_figures (accessed).
- [62] Litemax, "Digital Signage Display: 1968-U," 2021. [Online]. Available: <https://www.litemax.com/document-detail/0/1156/>.
- [63] Danphone. "Danphone Communication System: Flexible IP-based radio communication system for coast stations, ports and off-shore installations." <https://www.danphone.com/media/1510/danphone-communication-system.pdf> (accessed).

- [64] Teltonika. "RUT950." https://wiki.teltonika-networks.com/view/RUT950_MTBF (accessed.
- [65] MarineNav. "Leviathan 17i Fanless." https://www.marinenav.ca/pdf/comp/leviathan17i_fanless.pdf (accessed.
- [66] H. Xu, G. Li, and J. Liu, "Reliability analysis of an autonomous underwater vehicle using fault tree," in *2013 IEEE International Conference on Information and Automation (ICIA)*, 2013: IEEE, pp. 1165-1170.
- [67] E. Brocken, "Improving The Reliability Of Ship Machinery: A Step Towards Unmanned Shipping," 2016.