

Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
 Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Boudier, Roger Flage, Marja Ylönen
 ©2025 ESREL SRA-E 2025 Organizers. Published by Research Publishing, Singapore.
 doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P3941-cd

Leveraging Collision Avoidance Robustness to Establish Situational Awareness Requirements: A Closed Loop Simulator Approach

Grunde Løvoll

Group Research and Development, DNV, Norway. E-mail: grunde.lovoll@dnv.com

Kristian B. Karoliuss

Group Research and Development, DNV, Norway. E-mail: kristian.bertheussen.karoliuss@dnv.com

Henrik Stokland Berg

Group Research and Development, DNV, Norway. E-mail: henrik.stokland.berg@dnv.com

Peter Keenan Morris

Department of Electronic Systems, Faculty of Information Technology and Electrical Engineering, NTNU, Norway. E-mail: peter.k.morris@ntnu.no

Using a maritime traffic simulator to analyze the performance of autonomous navigation functions, specifically Collision Avoidance (CA), and how its performance depends on the quality of input from the Situational Awareness (SA) system is the target of this paper. To that end, a closed-loop simulator is developed that can take the SA “ground-truth” of a given maritime traffic situation, add modeled noise and other perturbations to ship positioning and world-states, feeding this as input to the CA system. For several scenarios, the quality (noise and error rate) of the input to the CA system is systematically varied, and performance is evaluated with respect to collision risk and compliance with a selected set of the International Regulations for Preventing Collisions at Sea (COLREG). This type of robustness assessment is crucial as it directly governs performance requirements for the SA system and its physical sensors used for sensing the environment and monitoring the condition of the ship. This paper provides examples of how quantifying the robustness of CA systems can determine performance criteria for SA systems.

Keywords: autonomy, assurance, safety, navigation, robustness, simulation testing, collision avoidance

1. Introduction

Recent advancements in artificial intelligence (AI), machine learning (ML), computer hardware and sensor technology have significantly increased interest and optimism in autonomous transportation. Impact on safety and logistical efficiency is expected to be substantial, particularly within the automotive industry (Eltoweissy et al. (2010); Burghout et al. (2015)). This optimism has spurred substantial investments in autonomous technology within the automotive sector (Deichmann et al. (2023)). Opportunities in the maritime domain are also being actively explored and developed for both research and commercial projects (Gu et al. (2021); Ziajka-Poznańska and Montewka (2021); de Vos et al. (2021); Skredder-berget (2025)). Human errors are believed to be the dominant cause of shipping accidents (Roth-

blum (2000); Porathe et al. (2018)), and the introduction of autonomous navigation systems and other advanced functions are anticipated to significantly reduce collision accidents.

However, humans excel at comprehending and managing complex dynamic situations and environments (Wróbel (2021)), making the complete replacement of human operators more challenging than initially anticipated. Despite recent advancements in autonomous technologies across the transportation industries (Waymo (2025); Apollo (2025); Skredderberget (2025)); the realization of fully autonomous transportation concepts remains elusive.

The technical challenge is twofold: first, the system has to perform equal to or better than human operators (IMO (2024)). Second, the technology provider, ideally with third-party verifica-

tion, has to demonstrate and document this performance. This is known as “the assurance challenge” and is present in both the automotive and shipping industries (Kalra and Paddock (2016); Koopman and Widen (2023)).

At the core of this challenge lies the fact that it is nearly impossible to get sufficient test coverage using only data from real operations. This problem is even more pronounced in the maritime industry, where ships and technical solutions are often tailor-made to individual scenarios rather than mass-produced.

The solution to this problem involves two key components: system modularization (DNV (2024)) and the use of simulators for sufficient test coverage (Huang et al. (2016); Koren and Kochenderfer (2019); Pedersen et al. (2020)). With a modular approach, high-level performance and safety goals can be identified and defined as measurable performance requirements for technical modules or complete systems. These requirements can then be propagated to underlying functions by studying how the performance of the high-level functions or systems depends on the quality and uncertainty of the output from lower-level functions.

This paper addresses the current lack of explicit assurance criteria by providing a concrete example of how to use system modularization and a simplified ship scenario simulator to analyze and measure the performance of safety-critical functions for autonomous navigation, specifically the Collision Avoidance (CA) system, and how its performance depends on the quality of input from the external Situational Awareness (SA) system.

The paper is organized as follows: Section 2 outlines and describes the system under study, the safety goals and requirements, methodology and simulator uses in the study. Section 3 presents the results, which are discussed in Section 4. Finally, Section 5 concludes and reflects on the way forward.

2. Methodology

2.1. Autonomous navigation

The system under study is an autonomous navigation system. The high-level components of this

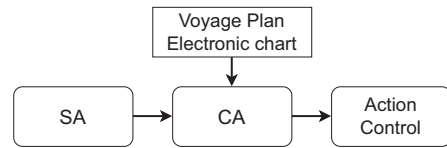


Fig. 1. Schematic high-level overview of the main functions and data flows of an autonomous navigation system.

system, along with rules and indicators for safe and compliant ship maneuvering, are described below.

The high-level functions of an autonomous navigation system (illustrated in Fig. 1) include: A SA system, which collects, analyzes and fuses data from various sensors (e.g. camera, lidar, radar, etc.) to generate a dynamic external world model and own-ships’s state. A CA system, which combines the output from the SA system with the voyage plan and electronic chart data to monitor collision risks. It provides a safe and COLREG-compliant (IMO (1972)) CA maneuver (i.e. a local alternation to the voyage plan) in case of identified risks. Finally, an action control system, which actuates and performs the maneuver provided by the CA system.

The SA system can be divided into internal and external components. The internal component monitors and assesses the internal state and capabilities of the Own-Ship (OS) and its systems, while the external system monitors the external navigation environment, and detects dynamic and static obstacles.

The output from the external SA system is a digital representation of the relevant physical environment (world model) used by the CA system to evaluate risks and generates safer alternative navigation routes. This paper focuses only on the external part of the SA system. For a more detailed description of the functions and data flows of autonomous navigation systems see Karoliuss et al. (2024).

The overall goal of the autonomous navigation system is to perform safe (COLREG-compliant) navigation, which manifests through the realization of a given traffic encounter scenario, and the

actual performed maneuver actions, the position, heading and speed, throughout the scenario in relation to the target-ships (TS). Given this data, safety performance and COLREG-compliance can be evaluated (Woerner et al. (2019)). Safety performance in traffic encounter scenarios have in this paper been evaluated using the distance between OS and TS at closest point of approach (DCPA)^a while the number of significant course changes have been used as an indication of COLREG performance, as few and apparent maneuver is specifically required in COLREG.

With safety evaluation criteria for the CA system established, it is possible to explore how the quality of the SA output (CA input) affects performance of the CA system in a simulated environment. This is done by running a limited set of predefined TS encounters (Pedersen et al. (2023)) multiple times with controlled noise and bias added to the output from the simulated SA system.

The simulator also models the physical control and maneuverability attributes of the simulated ships, and thus models the OS ability (or lack thereof) to accurately follow the CA system's control guidance.

2.2. Simulator framework

The results presented in this paper have been generated using an in-house ship encounter simulator. The simulator is a closed-loop simulator implemented in the Python programming language. The simulator does not include detailed spatial information of the simulation environment, and rather focuses on the ships' dynamics, and can therefore be considered a low-fidelity implementation. The high-level architecture of the simulator is illustrated in Fig. 2, and it contains multiple pluggable modules simulating the various components of an autonomous navigation system.

To generate complex ship encounter scenarios, the simulator is capable of simulating the dynamics of multiple ships simultaneously, each ship having its own (possible different) SA, CA and ac-

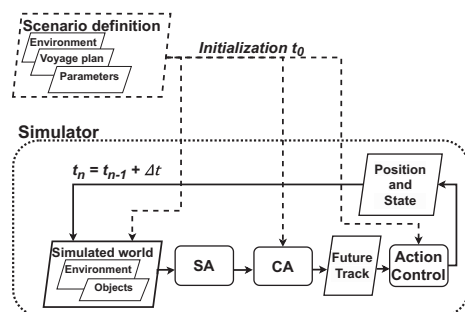


Fig. 2. Schematic overview of the autonomous navigation simulator for a single ship. The simulator is a closed loop where the initial state and planned voyage is fed into the simulator at t_0 . The world model is updated in each time step t_n , the simulated world is perceived by the SA module, the CA module assesses the situation and output desired future track/route which is implemented by the action control system. The output of each time step is the ship position and state (speed and heading) in the next time step. To create scenarios with multiple ships the simulator is capable of running one loop pr. ship in the scenario.

tion control systems. Running the simulation loop consists of the following actions, data flows and simulation modules: The scenario is initialized by importing the initial scenario, containing a voyage plan and the necessary ship parameters for all ships in the encounter scenario. The used format is the open-source Maritime Schema developed by Pedersen et al. (2023). Note that for this paper, the scenarios are currently limited to open, calm water, excluding bathymetrical and topographical features, as well as environmental factors such as wind, waves, and currents.

An external SA simulation model generates an estimated representation (world model) from the simulated ground truth data of the navigation environment around the OS. In this case the model includes the position, speed and heading of TS. In its simplest form, the external SA module reproduces the ground truth data in the reference frame of the ship in question (more details on the simulated SA system implemented in this paper is found in Sec. 2.3).

The CA system used is a simplified version of the Scenario-Based Model Predictive Control (SB-MPC) algorithm found in in Tengesdal and

^aCPA is position at Closest Point of Approach, DCPA is Distance between OS and TS at Closest Point of Approach

Johansen (2023).

The action control system contains two components: a speed and tracking control system implemented using an adaptive line-of-sight (ALOS) guidance algorithm (Johansen et al. (2016)) and a simplified PID controller (Fossen (2023)). Additionally, the physical properties of ship dynamics in three-degrees of freedom used is the ship maneuverability model for the ferry from the IEC track-control standard (IEC (2014)). The motion model is a simplified representation of ship motion but includes the most essential features such as reaction to machinery commands and propulsion forces, straight-line resistance, rudder commands and forces, and sideslip motion. For further details on ships specifications and maneuvering

characteristics for the ship models used in this study see Karolius et al. (2024).

The output in each pass of the simulation loop includes the updated position and state of each simulated ships. This updated ground truth is then fed into the simulated SA system in the next iteration. Fig. 3 illustrates a simulation of a single head-on traffic situation where the simulated OS makes a CA maneuver.

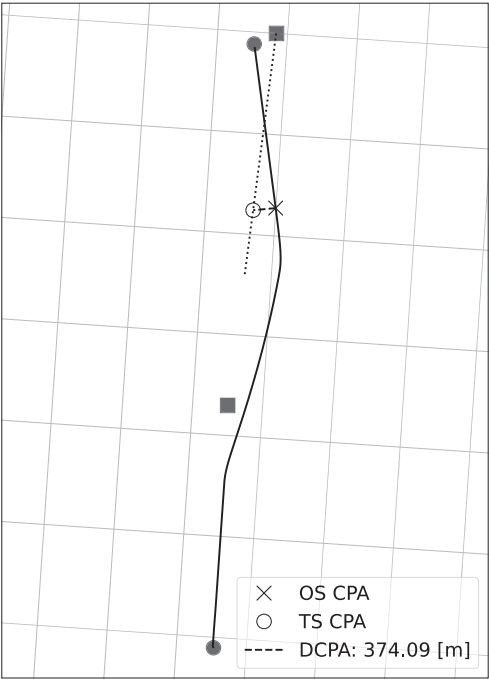


Fig. 3. Output from a single traffic situation, showing OS (with CA system) and TS in a head-on situation. The TS is following a straight line (moving top to bottom) and OS (moving bottom to top) is executing a CA maneuver. CPA and DCPA is indicated, and OS external SA is the ground truth. Initial voyage plan way points for OS and TS shown as solid circles and squares respectively. For reference the distance between the way points for OS is 10000 m.

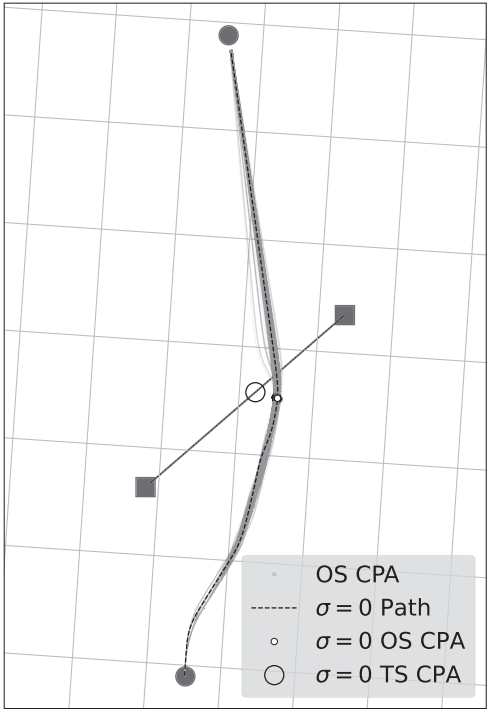


Fig. 4. Visualization of 1000 runs of the crossing scenario, OS moving bottom to top and TS crossing from upper right to lower left of figure. In each run random imposed detection noise with standard deviation (σ) 20 m for the position of the TS is added to SA system detection. The resulting trajectories of the OS is indicated as gray scale density map, and the corresponding situation without noise is depicted in the black dotted line with corresponding CPA. For reference the distance between the way points for OS is 10000 m.

2.3. External SA output quality and measuring CA safety performance

To better reflect real-world conditions, the simulator includes settings that allow for the introduction of configurable noise and bias into the simulated measurements. Adjusting these noise parameters allows for assessing the robustness of the CA in adapting to various scenarios and noise levels. High levels of noise may lead the OS to maneuver too close to the TS, and fluctuating noise can lead the CA to (needlessly) recalculate and follow a new course.

The simulator also features a tracker that models the expected behavior of TSs, smoothing over noisy measurements or predicting the trajectory of ships that have temporarily dropped out. This tracker can be adjusted to be more or less adaptive to sudden changes, impacting the decisions made by the CA.

Results are obtained through running a statistically significant number of simulations of a given traffic situation for each of a varied magnitude of standard deviation in the Gaussian noise of the detection measurements. An example of multiple realizations of the same traffic situation for a fixed noise level is shown in Fig. 4. The main measurements performed in the simulation experiments are the mean and standard deviation of DCPA ($\overline{DCPA}$ and SD) and mean number of significant course changes, with corresponding standard deviation, recorded over all simulations for each magnitude of Gaussian noise added to the detection measurements.

3. Results

This section outlines the results from the simulated crossing (Fig. 4) and head-on situations (Fig. 3). In both scenarios, the standard deviation in the imposed noise (σ) is systematically varied, and the $\overline{DCPA}$, the number of significant course changes, and the corresponding standard deviations are measured.

Fig. 5 and Fig. 6 presents the DCPA results for the head-on and crossing situations, respectively. In both cases, the $\overline{DCPA}$ values decrease (indicating a reduced safety margin) and the variation increases. This effect is most noticeable in the

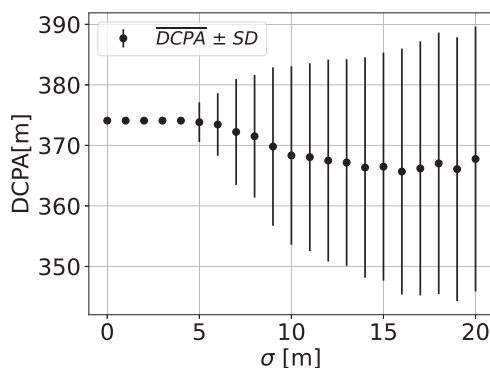


Fig. 5. Mean and standard deviation of the DCPA ($\overline{DCPA}$ and SD) between OS and TS as function of SA noise σ in a head-on situation. Each datapoint corresponds to 1000 simulations.

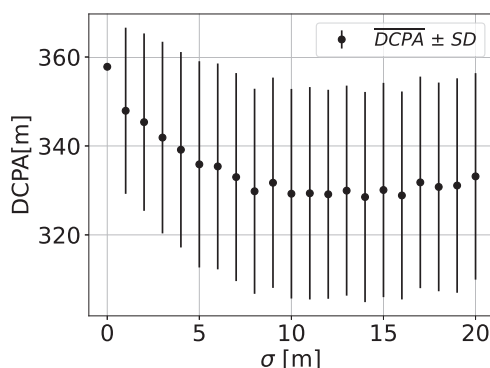


Fig. 6. Mean and standard deviation of the DCPA ($\overline{DCPA}$ and SD) between the OS and TS as function of SA noise σ in a crossing situation. Each data-point corresponds to 1000 simulations.

crossing scenario, where $\overline{DCPA}$ is reduced from 358 m (without SA noise) to ~ 330 m when $\sigma \geq 8$ m. The effect is also noticeable for low noise values in the crossing scenario. While head-on situations requires more noise ($\sigma \geq 5$ m) to be noticeable, and the overall effect appears smaller, with $\overline{DCPA}$ reduced from 374 m without SA noise to ~ 365 m for $\sigma > 15$ m. These findings indicate that the CA system's robustness to input noise is situation-dependent, highlighting the need for testing across a wide range of scenarios.

Fig. 7 and Fig. 8 show the mean number of course changes above 10° performed by the OS,

with corresponding standard deviation for the head-on and crossing situations, respectively^b. In both cases, the number of course changes increases, and the variation in realized values increases along with increased input SA noise. The number of executed significant course changes serves as a (crude) measure of COLREG performance. The trend is similar for the number

^bA course change exceeding 10° is established as a somewhat arbitrary threshold for an “apparent maneuver” in accordance with COLREG.

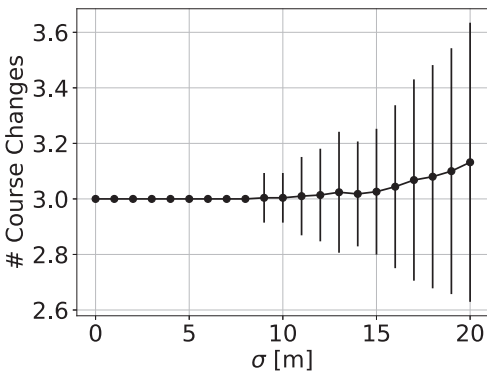


Fig. 7. Mean and Standard deviation of number of course changes above 10° for the OS as function of SA noise σ in a head-on situation. Each data point corresponds to 1000 simulations.

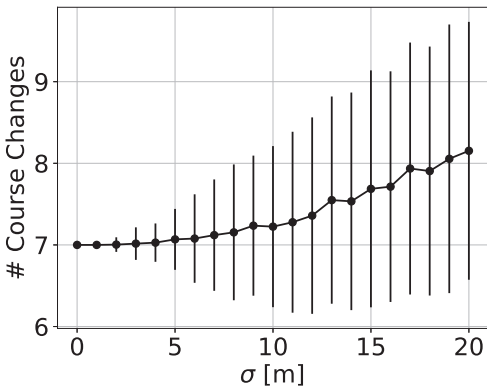


Fig. 8. Mean and Standard deviation of number of course changes above 10° for the OS as function of SA noise σ in a crossing situation. Each data-point corresponds to 1000 simulations.

of course changes as for DCPA; the CA system appears less sensitive to input noise in the head-on situation than for the crossing situation. In the head-on case (Fig. 7), the average number of course changes and the variation increases for $\sigma > 9$ m. For the crossing situation (Fig. 8), the overall number of course changes is higher even without input SA noise (~ 7 vs. ~ 3), likely reflecting a weakness of the chosen CA algorithm. The crossing situation also shows a more pronounced response, both in the average number of course changes and standard deviation, when SA noise is introduced.

4. Discussion

The results and study presented in this paper have several significant limitations. The results are based on a limited number of traffic scenarios and OS types, and the vast “ground truth scenario space” has not been fully explored. The simulator used is a relatively simple low-fidelity implementation that does not account for external environmental factors. The CA system tested does not include grounding avoidance, and all the simulations are conducted in open water, excluding bathymetry and topography. The method used to evaluate safety performance, based on mean and variation in DCPA and number of significant course changes, is simplistic and does not include an extensive evaluation of “COLREG performance”. The method for simulating errors and uncertainty in the SA system is also simplistic and not based on measurements or statistics from the underlying tracker, sensor, or sensor fusion systems.

The configuration of the CA system significantly impacts the results. More frequent CA updates increase sensitivity to detection changes, while increased safety margins can reduce system efficiency and impact course changes. More advanced tracking systems can mitigate the effects of detection noise fluctuations but may become more susceptible to inaccuracies in initial measurements. Future work should focus on measuring how the CA prioritizes various COLREGs, safety targets, and efficiency goals, although determining appropriate metrics will be challenging.

Despite these shortcomings, the presented concept and method are considered novel and could be further developed to ensure the overall safety of autonomous navigation systems. The modular design of the simulator allows for the addition and modification of functional modules^c, and can be integrated with automatic generation of traffic situations for systematic exploration of the “scenario space” and smart testing, as suggested by Pedersen et al. (2020). These aspects, combined with the efficiency of a low-fidelity simulator, enable substantial expansion of the test space and the development of scalable, cost efficient, and automated tools for testing autonomous navigation systems.

Future work will focus on expanding and implementing more features in the simulator to establish concrete requirements for autonomous ship system components and develop tools for third-party verification and testing. In the near future, the ambition is to include more OS ship types in the tests and to add support for more realistic external SA systems, including tracker (Vasstein et al. (2023)), simulated detections (Vasstein et al. (2020)) and sensor fusion. Additionally, a CA system with grounding avoidance functionality and electronic chart data should be included. Another interesting avenue would be to investigate typical sensor failure modes and portray accurate models of their expected noise profiles to see the impact on tracker and CA performance. Finally, presenting more novel situations for the CA to deal with, including failure by TS to adhere to COLREGs, or malicious GPS jamming, etc. would be vital in validating CA systems for real-world use.

5. Concluding remarks

Combining a modular system architecture with real-life testing and extensive simulator-based testing at different fidelity levels appears to be the only viable path towards achieving sufficient test coverage and documenting the safety of autonomous navigation systems.

Extensive simulator-based testing is particu-

larly important in the maritime sector, where the number of ships with such systems is low, and the concepts under development vary greatly in terms of installed hardware and operational design domains.

The concept and method presented in this paper offer a promising approach to testing and ensuring that high-level safety and performance goals are adequately supported by the functional components of the installed system.

Acknowledgement

The work outlined in this paper was produced through the research project SAFE Maritime Autonomous TEchnology (SAFEMATE – grant number: 327903) and SFI AutoShip project, both supported by the Research Council of Norway.

We would also acknowledge Adrian Pavlak's and Simen Rørvik's contributions to the initial development of the simulator utilized in this study.

References

- Appollo (2025). Apollo intelligent vehicle. <https://en.apollo.auto/apollo-self-driving>. Accessed: 2025-01-09.
- Burghout, W., P.-J. J. Rigole, and I. Andréasson (2015, January). Impacts of shared autonomous taxis in a metropolitan area. In *The Proceedings of the 94th Annual Meeting of the Transportation Research Board*, Washington DC.
- de Vos, J., R. G. Hekkenberg, and O. A. Valdez Banda (2021). The impact of autonomous ships on safety at sea – a statistical analysis. *Reliability Engineering & System Safety* 210, 107558.
- Deichmann, J., E. Ebel, K. Heineke, R. Heuss, M. Kellner, and F. S. M. . Company) (2023). Autonomous driving's future: Convenient and connected. <https://www.mckinsey.com/>. Accessed: 2025-01-06.
- DNV (2024). Class guideline dnv-cg-0264 autonomous and remotely operated vessels. <https://standards.dnv.com/explorer/document/61F34A49A17F4929A6BE38765C68A19B/5>. Accessed: 2025-01-02.
- Eltoweissy, M., S. Olariu, and M. Younis (2010). Towards autonomous vehicular clouds. In J. Zheng, D. Simplot-Ryl, and V. C. M. Leung (Eds.), *Ad Hoc Networks*, Berlin, Heidelberg, pp. 1–16. Springer Berlin Heidelberg.
- Fossen, T. I. (2023). An adaptive line-of-sight (alos) guidance law for path following of aircraft and ma-

^cIncluding commercial autonomous navigation systems components.

- rine craft. *IEEE Transactions on Control Systems Technology* 31(6), 2887–2894.
- Gu, Y., J. C. Goetz, M. Guajardo, and S. W. Wallace (2021). Autonomous vessels: state of the art and potential opportunities in logistics. *International Transactions in Operational Research* 28(4), 1706–1739.
- Huang, W., K. Wang, Y. Lv, and F. Zhu (2016). Autonomous vehicles testing methods review. In *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, pp. 163–168.
- IEC (2014, February). Iec 62065:2014 ed2 maritime navigation and radiocommunication equipment and systems - track control systems - operational and performance requirements, methods of testing and required test results, 2nd edition. Technical Report ISBN 978-2-8322-4858-4, International Electrotechnical Commission (IEC).
- IMO (1972). Convention on the international regulations for preventing collisions at sea (COLREG). <https://www.imo.org/en/About/Conventions/Pages/COLREG.aspx>. Accessed: 2025-01-02.
- IMO (2024). Imo goal-based standards. <https://www.imo.org/en/OurWork/Safety/Pages/Goal-BasedStandards.aspx>. Accessed: 2025-01-02.
- Johansen, T. A., T. Perez, and A. Cristofaro (2016). Ship collision avoidance and colregs compliance using simulation-based control behavior selection with predictive hazard assessment. *IEEE Transactions on Intelligent Transportation Systems* 17(12), 3407–3422.
- Kalra, N. and S. M. Paddock (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A: Policy and Practice* 94, 182–193.
- Karoliuss, K. B., G. Løvoll, J. Cichowicz, Ø. Engelhardttsen, and A. Jørgensen (2024, October). Navigating the challenges of assuring safe autonomous navigation. In *Proceedings of the 2nd International Conference on the Stability and Safety of Ships and Ocean Vehicles, Wuxi China*.
- Karoliuss, K. B., G. Løvoll, T. A. Pedersen, A. L. Pavlak, and S. A. Rørvik (2024). Rina autonomous ships conference. In *2024 RINA Conference proceedings*.
- Koopman, P. and W. H. Widen (2023). Breaking the tyranny of net risk metrics for automated vehicle safety. *University of Miami Legal Studies Research Paper November 15*.
- Koren, M. and M. J. Kochenderfer (2019). Efficient autonomy validation in simulation with adaptive stress testing. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 4178–4183.
- Pedersen, T. A., J. A. Glomsrud, E.-L. Ruud, A. Simonsen, J. Sandrib, and B.-O. H. Eriksen (2020). Towards simulation-based verification of autonomous navigation systems. *Safety Science* 129, 104799.
- Pedersen, T. A., C. Vasanathan, K. Karoliuss, Ø. Engelhardttsen, K. P. Houweling, and A. Jørgensen (2023). Generating structured set of encounters for verifying automated collision and grounding avoidance systems. In *Journal of Physics: Conference Series*, Volume 2618, pp. 012013. IOP Publishing.
- Porathe, T., Å. Hoem, Ø. Rødseth, K. Fjortoft, and S. O. Johnsen (2018). At least as safe as manned shipping? autonomous shipping, safety and “human error”. In *Safety and Reliability–Safe Societies in a Changing World*, pp. 417–425. CRC Press.
- Rothblum, A. M. (2000). Human error and marine safety. In *National Safety Council Congress and Expo, Orlando, FL*, Volume 7.
- Skredderberget, A. (2025). The first ever zero emission, autonomous ship. <https://www.yara.com/knowledge-grows/game-changer-for-the-environment/>. Accessed: 2025-01-06.
- Tengesdal, T. and T. A. Johansen (2023). Simulation framework and software environment for evaluating automatic ship collision avoidance algorithms*. In *2023 IEEE Conference on Control Technology and Applications (CCTA)*, pp. 186–193.
- Vasstein, K., E. F. Brekke, R. Mester, and E. Eide (2020, nov). Autoferry gemini: a real-time simulation platform for electromagnetic radiation sensors on autonomous ships. *IOP Conference Series: Materials Science and Engineering* 929(1), 012032.
- Vasstein, K., Ø. K. Helgesen, and E. F. Brekke (2023). Hellinger metrics for validating high fidelity simulators using target tracking. In *Modelling and Simulation for Autonomous Systems*, Cham, pp. 25–43. Springer International Publishing.
- Waymo (2025). Waymo driver. <https://waymo.com/waymo-driver/>. Accessed: 2025-01-09.
- Woerner, K., M. R. Benjamin, M. Novitzky, and J. J. Leonard (2019). Quantifying protocol evaluation for autonomous collision avoidance: Toward establishing colregs compliance metrics. *Autonomous Robots* 43, 967–991.
- Wróbel, K. (2021). Searching for the origins of the myth: 80% human error impact on maritime safety. *Reliability Engineering & System Safety* 216, 107942.
- Ziajka-Poznańska, E. and J. Montewka (2021, 05). Costs and benefits of autonomous shipping—a literature review. *Applied Sciences* 11, 4553.