Chapter 9

Toward a structured framework for control performance and safety assessment for Maritime Autonomous Surface Ships

V. Garofano[1], Y. Pang[1] and R.R. Negenborn[1]

As the maritime industry moves toward fully autonomous operations, it is becoming increasingly important to assess the control performance and safety of Maritime Autonomous Surface Ships. This chapter presents a structured framework designed to facilitate testing and data collection using autonomous ship systems, thereby supporting the verification that autonomous operations align with International Maritime Organization standards. We discuss the integration of key hardware and software components for robust autonomous operation and evaluate these systems using analytical performance criteria in both simulated and real-world scenarios. The chapter concludes by proposing new key performance indicators necessary for the continued development of autonomous maritime systems. Through extensive datasets and collaborative research via Open Science-focused algorithms and designs, we aim to set the groundwork for future advancements in this field.

9.1 Introduction

The technological advancement in autonomous navigation technologies has initiated a significant shift across various industries, with the maritime sector exploring the advancement toward fully autonomous Maritime Autonomous Surface Ships (MASSs). This technology promises enhanced efficiency, safety, and sustainability in maritime operations, yet it simultaneously poses complex challenges in terms of control performance and safety assessment evaluation [2,26]. Traditionally, these aspects have been governed by standards set by the International Maritime Organization (IMO), which, while comprehensive, were not designed with autonomous technologies in mind [7]. Therefore, it starts to be important to rethink these standards to promote the integration of smart navigation systems within the regulatory frameworks that ensure maritime safety and can evaluate the performance of decision-making algorithms [1].

[1]Department Maritime and Transport Technology, Faculty Mechanical Engineering, Delft University of Technology, The Netherlands

MASSs operate under different circumstances than manned ships, requiring robust control algorithms and hardware systems that can perform reliably in changing and often unpredictable marine environments [4,24]. To address these needs, this chapter proposes a structured test framework that can perform tests following existing IMO standards with state-of-the-art autonomous navigation technologies [20,25]. This approach is aimed at redefining control performance and safety assessment methodologies to accommodate the autonomous capabilities of future vessels [6,8].

Different policies and guidelines are already in place to regulate the safety of ship navigation. The SOLAS (Safety of Life at Sea) Convention, for example, regulates specific equipment and procedures to ensure the safety of navigation, fire safety systems, life-saving appliances, and communication protocols. However, in the context of autonomous ships, these provisions necessitate reinterpretation [5,11]. Autonomous vessels must be equipped with advanced navigation systems that not only comply with these regulations but also enhance them through technology. For example, automated systems and life-saving protocols need redesigning to function effectively with less human intervention on board [10]. Under the COLREGs (International Regulations for Preventing Collisions at Sea), the ability of ships to avoid collisions autonomously and safely becomes paramount [9]. Traditional rules designed for human interpretation must now be encoded within the upcoming AI systems of autonomous ships. These vessels must be capable of interpreting complex maritime scenarios and making real-time decisions to maintain navigational safety [21,22]. This requires a sophisticated integration of sensor data and decision-making algorithms that can robustly interpret and act according to the COLREGs [3,21].

The MARPOL (International Convention for the Prevention of Pollution from Ships) convention focuses on minimizing environmental impact of maritime operations. Autonomous ships offer a promising path to enhanced compliance with these regulations. Through optimized route planning and feedback controls, these ships can reduce fuel consumption and emissions, significantly lowering their environmental footprint [19].

Although the STCW (Standards of Training, Certification, and Watchkeeping) convention primarily addresses crew qualifications, its adaptation for autonomous shipping is becoming important as well. Operators who remotely monitor and control autonomous fleets must receive specialized training that reflects the technical and operational scenarios of unmanned shipping. This ensures that even without a traditional crew, the vessels are operated safely and effectively under rigorous oversight standards [36].

Performance standards for electronic navigational aids, such as ECDIS, AIS, and radar, are outlined by the IMO to ensure that maritime traffic is managed safely and efficiently [25]. For autonomous ships, these systems must also integrate advanced technological solutions that enhance their functionality and reliability [31]. This involves the implementation of redundant systems and robust sensor fusion technologies that can deliver precise and reliable navigation data in real time [12].

This chapter will explain and analyze the Researchlab Autonomous Shipping (RAS) idea of a structured framework, describing the hardware and software components that constitute autonomous ship systems as a whole. From sensors and communication devices to complex algorithms responsible for decision-making and obstacle avoidance, the framework will cover the spectrum of technological requisites and assessment performance criteria for autonomous operations [35]. This discussion will highlight the technological support of autonomous ships and investigate the reliability, redundancy, and real-time processing needs that these systems demand [18,33].

The practical application of these technologies will be described as well, where various configurations and scenarios, including open water tests and towing tank experiments, are employed to refine and validate the control strategies of autonomous ships [16]. The chapter will also propose the development of new Key Performance Indicators (KPIs) necessary for assessing the efficacy and safety of autonomous navigation systems [14,19,29].

The main contribution of the RAS-structured framework is its innovative approach to integrating and evaluating autonomous navigation technologies within existing maritime safety and performance standards. This framework is designed to bridge the gap between traditional maritime regulations and the advanced capabilities of MASSs. It redefines control performance and safety assessment methodologies, ensuring that autonomous vessels can operate reliably and safely under various conditions [5,38]. The RAS framework systematically incorporates robust control algorithms, advanced navigation systems, and sophisticated sensor fusion technologies, ensuring that these systems not only comply with but also enhance current regulatory requirements [13,17,28]. Additionally, it emphasizes the importance of developing new KPIs and sharing extensive datasets to propel the advancement of autonomous maritime operations [24]. In this context, a structured framework refers to a comprehensive and methodical approach that investigates all critical aspects of autonomous ship systems, from hardware and software components to performance criteria and practical testing scenarios [23]. It ensures a cohesive integration of technologies, regulatory compliance, and performance evaluation, facilitating the reliable and safe development of autonomous maritime vessels [32,34].

9.2 The RAS framework

The RAS framework is designed to address these evolving needs by incorporating technologies and methodologies to enhance safety, performance, and compliance in autonomous maritime operations. By focusing on real-time data acquisition, advanced sensor capabilities, and robust communication systems, the RAS framework aims to provide a comprehensive test field that aligns with and potentially extends the test scenario in existing IMO standards.

Through the RAS framework, it is possible to design and validate collision-avoidance applications, enhance redundancy and reliability in navigation systems, and potentially increase environmental compliance, all while adapting safety equipment and procedures to fit the unmanned nature of autonomous ships. These advancements represent a step forward in the maritime research field, enabling possibilities for a safer, more efficient, and environmentally responsible future for autonomous maritime operations.

Central to the RAS framework is its focus on developing a reliable and robust communication layer, based on the Robotic Operating System (ROS) [30]. Traditional maritime operations, reliant on onboard human decision-making, face inherent delays in data processing and action implementation. On the contrary, the RAS framework emphasizes real-time, low-latency communication between autonomous ships and shore-based control centers. This strategy is designed to monitor decision-making algorithms' efficiency and provide a better insight into reaction times to navigational and operational changes.

To achieve this, the framework incorporates:

- **Real-time diagnostics**: Automated diagnostic tools continuously assess the health and functionality of onboard systems, with the possibility to flag any irregularities for immediate attention.
- **Redundancy management**: The framework can potentially operate multiple layers of redundancy for hardware and sensor systems, ensuring that a single point of failure does not compromise the vessel's operational capability or safety.
- **Dynamic data sharing**: Information from individual ships is shared into a centralized database accessible via a secure VPN, allowing for dynamic data sharing and enhanced collective decision-making algorithms.
- **Fleet-wide operational optimization**: Utilizing data from across the fleet, the system can identify patterns, predict potential issues, and optimize fleet operations through coordinated control strategies. This is especially important in complex scenarios such as navigating busy shipping lanes or conducting synchronized logistical operations.
- **Scalability and flexibility**: The networked approach is designed to be scalable, accommodating an increasing number of autonomous vessels without compromising performance or security. It also provides the flexibility needed to customize control strategies for different types of vessels and operational tasks.

These technologies rely on robust, high-speed communication infrastructures that enable the real-time transmission of vast amounts of data to and from the autonomous ships. By leveraging wireless data exchange and/or advanced satellite communications, the system ensures that data flows are not disrupted, which is important for maintaining operational integrity and safety. Figure 9.1 provides a schematic of the RAS framework.

To ensure reliable monitoring of the autonomous ship's dynamic parameters and navigation information, the vessel is equipped with specific sensors. These sensors are connected to a central CPU, which processes and forwards vessels' information to a shore monitoring and control center. The ships used in the RAS framework are equipped with the following sensors:

- **GNSS receiver:** This sensor receives feedback about the real-time positioning of the Autonomous Surface Vessel. To enhance accuracy, Differential GNSS technology is employed, which mitigates atmospheric errors with the help of a ground station. Another method to improve position feedback is the implementation of Real-Time Kinematic technology and mathematical-based filters such as the Kalman Filter, which help in refining the accuracy of the positioning data.

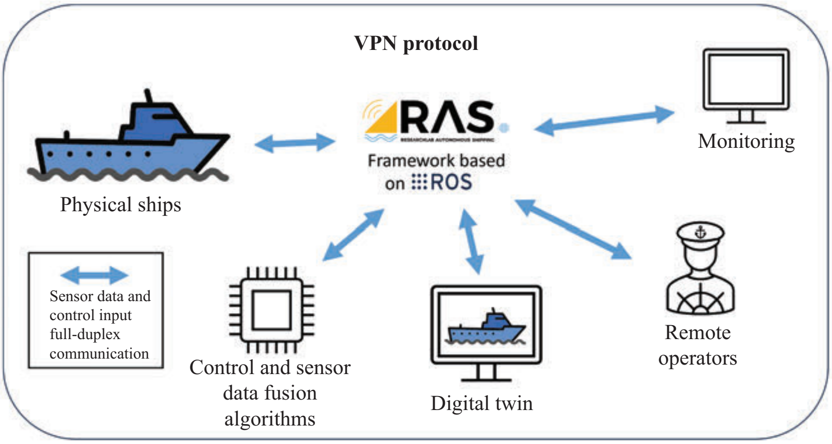

Figure 9.1 RAS framework overview

- **Stereo vision cameras:** Essential for autonomous ships to perceive their surroundings independently, these cameras are installed at various points around the ship to create a 360-degree view of the environment. The stereo vision technology not only helps in distance measurement but is also integrated with state-of-the-art AI computer vision technologies to autonomously detect and distinguish between different objects in the environment.
- **Inertial measurement unit (IMU):** Comprising a 6 Degrees of Freedom (DoF) accelerometer, gyro, and magnetometer, this sensor provides real-time data on various states of the ship. It helps in measuring forces applied to the ship's hull that can be used to extract hydrodynamic resistance values, angular velocity, and heading – crucial for understanding the internal state conditions without human onboard presence.
- **Encoders on propulsion shaft:** These measure the RPMs at which the ship's propellers are rotating, providing critical data for evaluating the thrust effort required by the control algorithm for precise navigation along a predefined path.
- **Energy measurement unit:** It evaluates the amount of mechanical and electrical energies required by the control algorithm under various navigational scenarios. This data is vital for developing control algorithms that optimize energy use efficiently.
- **Propeller azimuth angle feedback:** Whether using rudders or azimuth rotative propellers, this sensor provides feedback on the direction in which the propellers are pushing thrust, enhancing the capacity to assess control performance accurately.

All these sensors and their related data are then used in the assessment of control performance and safety of the autonomous navigation scenario. The combined data from these sensors are indeed important for determining whether the control algorithms are performing as expected.

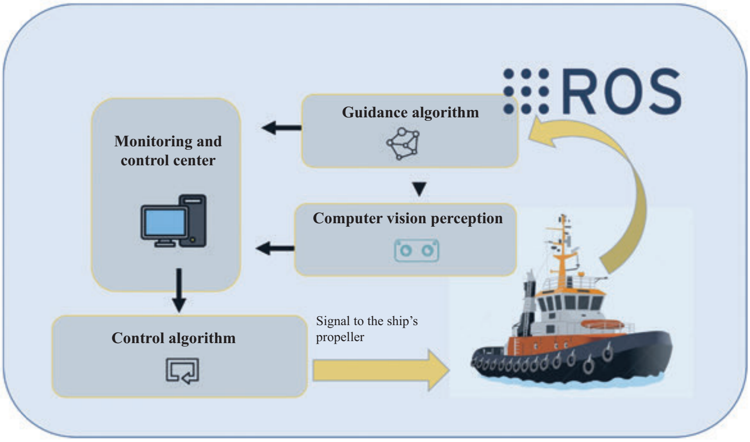

Figure 9.2 Overview of RAS framework control and feedback loop

The RAS framework provides control feedback closed-loop interaction between several hardware and software components, as shown in Figure 9.2 that can be used to check standard requirements in autonomous maritime operations, aligning with existing IMO standards and extending their scope to incorporate advanced technological capabilities [12].

Building on the general framework design described previously, we utilize this basis to gather data and conduct laboratory tests on autonomous ship systems. To ensure the reliability, safety, and effectiveness of control algorithms that manage the ship's automated systems, the testing involves a series of simulations and experiments in a controlled laboratory environment. This approach allows for methodical progress toward replicating more complex scenarios encountered in open water environments, thereby closely emulating real-world conditions, Figure 9.3.

The development of autonomous ship systems starts in simulation environments, where specialized tools try to recreate realistic maritime conditions. These digital testbeds must replicate the physical and operational dynamics accurately, ensuring that the simulations are as close to real-world scenarios as possible. This fidelity is achieved through state-of-the-art analytical models of aerodynamic and hydrodynamic interactions within the ship system, incorporating environmental disturbances such as currents, waves, and winds based on the latest research results.

Beyond virtual simulations, Hardware-in-the-Loop (HIL) testing incorporates physical components – such as sensors and actuators – into a simulated environment. This approach evaluates the hardware system in its entirety under controlled yet realistic conditions, bridging the gap between theoretical models and practical operations. Similarly, scale-model testing in facilities such as towing tanks or flume tanks (see Figure 9.4) provides critical insights into the physical dynamics of ships. These experimental setups yield valuable data on how well a vessel and its systems handle stability and maneuverability, thereby informing more robust design and validation processes.

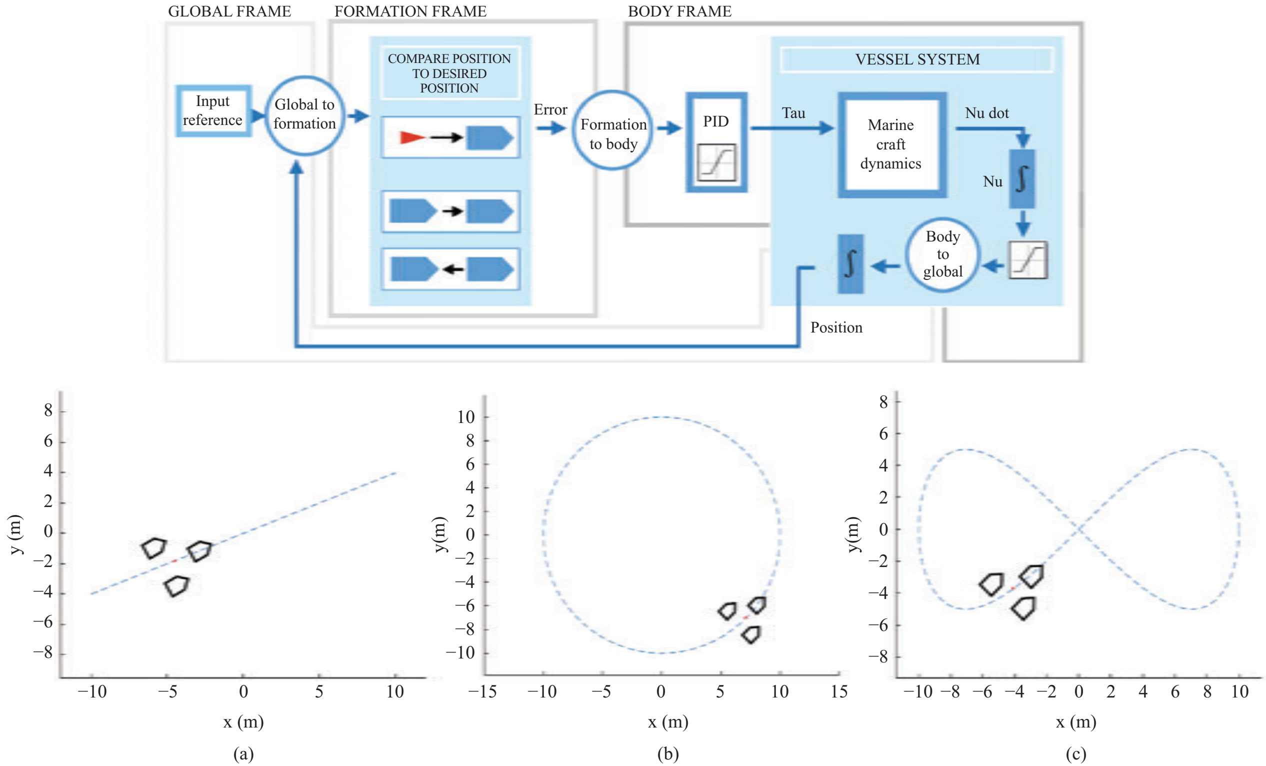

Figure 9.3 Simulation environments of ResearchLab Autonomous Shipping. (a) Straight line path (b) Circular path and (c) Lemniscate path.

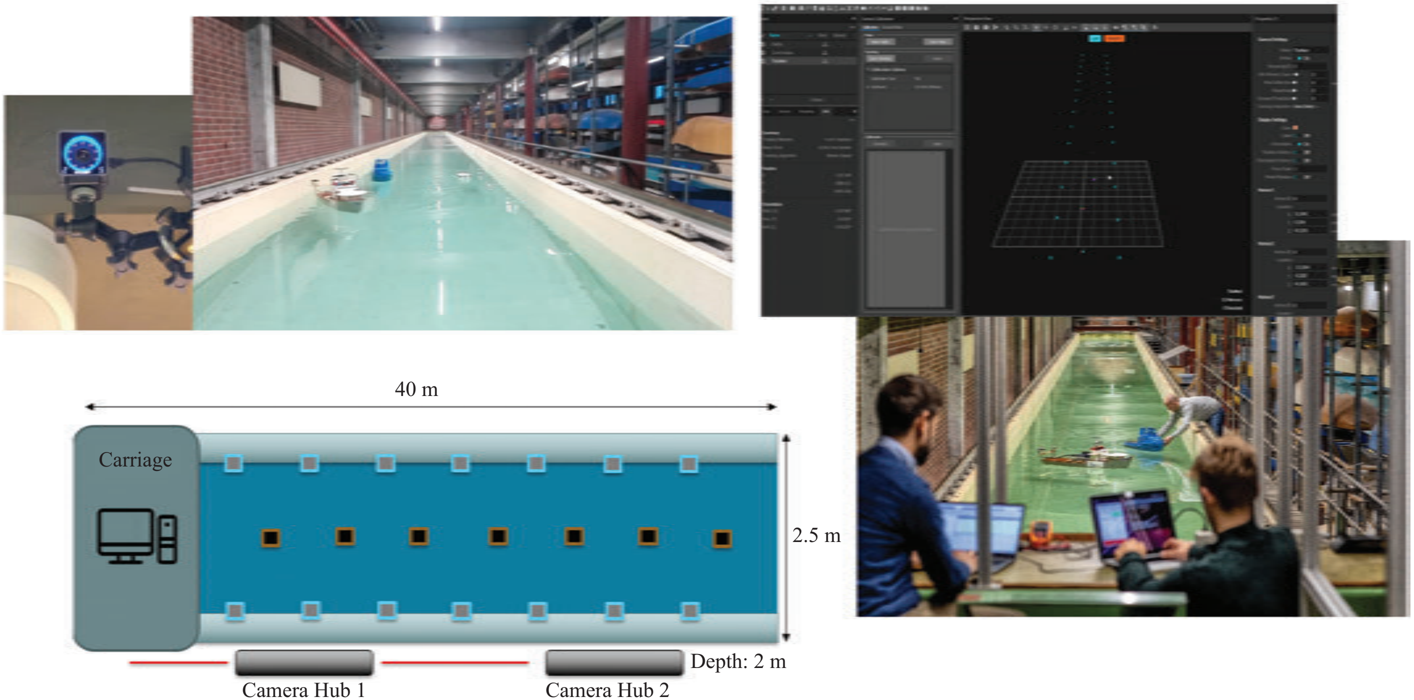

Figure 9.4 Experiments in water tank facilities at RAS

Figure 9.5 Experiments in Open Water Facilities at RAS

Throughout these tests, the real-time collection and analysis of data are essential. Monitoring the output from various sensors and systems during tests helps in fine-tuning the technology, enhancing the algorithms, and ensuring the systems' reliability across different scenarios. This iterative cycle of testing, data gathering, and analysis helps advance the systems toward real-world readiness (see Figure 9.5).

Moreover, ensuring that subsystems within an autonomous ship function in unison is another key aspect of experimental campaigns. Integration and interoperability testing in a laboratory setting can reveal potential communication and functionality issues, supporting robust operation across the vessel's navigation, communication, and operational systems.

Autonomous systems are then evaluated, Figures 9.6 and 9.7 provide examples to ensure compliance with international performance and safety standards, such as those outlined by the IMO.

Finally, the collaborative nature of research and development in this field drives innovation and refines autonomous systems. Partnerships among academia, industry leaders, and regulatory bodies are essential, as they consolidate expertise, resources, and perspectives. This collective effort not only fosters technological progress but also ensures alignment with global standards and practical maritime needs [15] This approach can support transitioning from theory to practice, turning innovative research outcomes into operational realities.

9.3 Assessment of the control and safety performance

Assessing the control and safety performance of the MASS is important for ensuring their reliable and safe autonomous operation. The methods for testing and assessing the performance of MASSs have been developed under the RAS framework [35], with relevant research projects contributing significantly to this field [29]. This section outlines the test scenario and the criteria that can be used for assessing the control

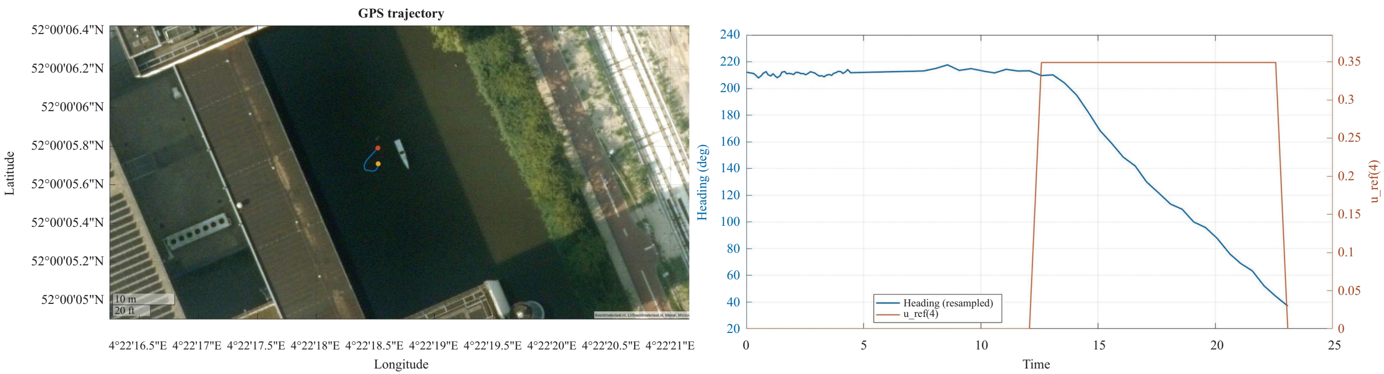

Figure 9.6 Example of maneuvering test and data collection for performance and operational safety assessment

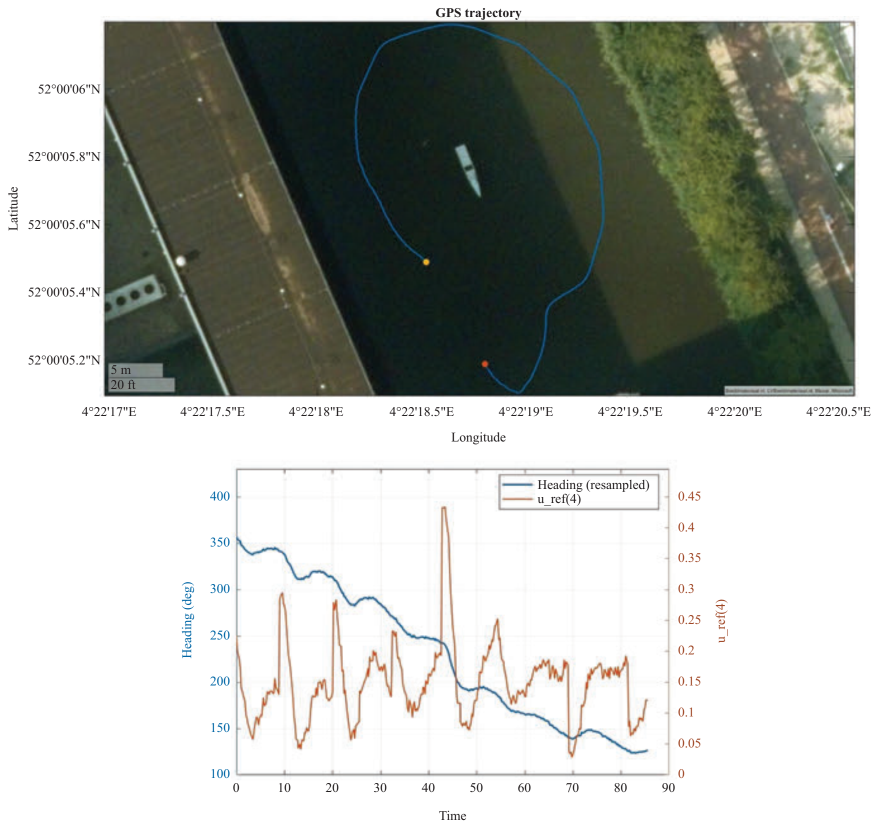

Figure 9.7 Autonomous path following dataquery

and safety performance of MASSs. Each scenario focuses on a different aspect of operation and performance, providing a comprehensive evaluation that covers safety, navigability, stability, and efficiency. The aim is to establish a set of criteria that can evaluate the vessel's ability to operate reliably and safely under various conditions.

By applying scenario-based testing to these main performance aspects, the idea is to have an objective evaluation of the MASS's performance. This methodical approach can evaluate whether an autonomous vessel meets required standards and performs effectively.

This section describes the primary testing scenarios, which include maneuvering, path following, collision avoidance, and multiple vessel coordination.

Maneuvering tests are designed to evaluate the MASS's ability to perform various movements such as turning, reversing, stopping, and changing speeds, focusing on the vessel's responsiveness and its capability in handling complex maneuvers safely and efficiently. Path-following tests examine the MASS precision in adhering to a predetermined course, testing its navigational systems under varying environmental disturbances to check accurate course maintenance and control. In collision avoidance tests, the MASS encounters both static and dynamic obstacles, testing its sensor systems and algorithms for effective detection and evasion, important for operational safety and reliability. Finally, multiple vessel coordination involves deploying several MASSs simultaneously to assess their collective operational capabilities, measuring their coordination capability, formation keeping, and cooperative behavior.

9.3.1 Maneuverability assessment aspects

9.3.1.1 Navigate in a straight line

Testing the autonomous surface ship for its ability to navigate straight is key for assessing navigation systems and stability. This is initially done in a controlled waterway to measure trajectory accuracy. This test, critical for ensuring efficient and safe operation, uses GNSS or camera systems to monitor deviations from a straight path at regular intervals Δt. Performance is measured by the average and maximum deviations from the intended path.

$$d(t) = \frac{|(x_{\text{end}} - x_{\text{start}})(y_{\text{start}} - y_{\text{mean}}(t)) - (x_{\text{start}} - x_{\text{mean}}(t))(y_{\text{end}} - y_{\text{start}})|}{\sqrt{(x_{\text{end}} - x_{\text{start}})^2 + (y_{\text{end}} - y_{\text{start}})^2}}$$

The MASS's performance is considered successful if the maximum and average deviations meet the defined acceptance thresholds, as described in Tables 9.1 and 9.2.

Table 9.1 Acceptance criteria for navigate in a straight line scenario

KPI	**Acceptance criteria**
Maximum deviation $\max\{d(t)\}$	Less than a_1 meters
Average deviation $\frac{\sum_{t=0}^{n} d(t)}{n}$	Less than a_2 meters

Table 9.2 Validation criteria for the mooring test

KPI	Acceptance criteria
Distance d	$d < a$ meters

9.3.1.2 The mooring test

The mooring test evaluates how precisely an autonomous surface vessel can use its stopping mechanisms to moor at a predetermined spot, crucial in crowded urban ports to avoid collisions and ensure safety. The test measures the MASS's ability to accurately stop at a specific berthing point, assessing navigational accuracy, control systems, and responsiveness. Effective docking reduces unnecessary maneuvers, saving time and energy, thus improving cost-efficiency.

A designated point, Ms(x_s, y_s), marks the target mooring location. The objective is for the ASV to halt as close to Ms as possible. The key performance indicator is the distance d between the actual stop point Mm(x_m, y_m) and Ms, calculated via GPS or cameras. The distance is given by:

$$d = \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2}$$

To determine the success of the mooring test, an acceptance value a is predefined based on the operational requirements of the MASS. The test is considered successful if the calculated distance d is less than the acceptance value a.

9.3.1.3 Maximum speed

The maximum speed test assesses the performance and energy usage of an autonomous vessel at its speed limit. It requires a long waterway where the ship can accelerate to its maximum speed, testing its ability to reach and maintain this speed without further acceleration, signifying a steady state.

This test evaluates the vessel's top speed and its effects on maneuverability, including turning radius and stability. High-speed tests also identify potential control system issues that may emerge only at high speeds.

Speed is measured with accelerometers and the vessel's position via GNSS or cameras. The primary KPI is the maximum speed, V_m. The performance meets expectations if V_m is within an acceptable range of the designed maximum speed V_d, denoted as a.

The collected data includes time points, coordinates, and calculated speeds:

The scenario is considered successful if

$$V_d - \max\{V(t)\} < a$$

where $\max\{V(t)\}$ represents the highest speed recorded during the test, and a is the predefined acceptance value.

9.3.1.4 Stopping test

The stopping test evaluates the ability of a MASS to halt using a stop engine-full astern maneuver. This maneuver determines the MASS's maximum stopping distance, a key safety metric. The test applies to autonomous vessels that can reverse and requires a long waterway for execution.

The test procedure starts as the MASS cruises at a set speed and then shifts to full astern to decelerate. The key performance indicators are the head reach (HR) and the track reach (TR). HR is the distance the MASS travels forward from deceleration onset to a full stop. TR is the total distance traveled until it completely stops.

$$HR = |x(t_1) - x(t_0)|$$

$$TR = \sum_{i=1}^{ts} \sqrt{(x(t_i) - x(t_{i-1}))^2 + (y(t_i) - y(t_{i-1}))^2} < 10L$$

Both HR and TR can be measured using GNSS, cameras, and distance meters. The performance of the test is considered satisfactory if the TR is smaller than 10 times the vessel length (L), with decreasing HR values indicating better performance. The test should be repeated at different speeds to comprehensively assess the MASS's stopping ability across its entire operational speed range.

9.3.1.5 Turning test

The turning test assesses the autonomous vessel capability to perform a U-turn, crucial for route changes or dock departures in various maritime settings. This maneuver tests the MASS's minimum turning radius at different speeds, key for effective maneuvering in tight waterways. The procedure starts with the MASS on a straight path, then the rudder is set to its maximum design angle, usually up to 35 degrees, to initiate a full circle turn to both the starboard and the port.

Measurements begin at the maneuver command (t_0) and end once the MASS completes the turn. The KPIs are the advance (Ad) and the tactical diameter (TD). These are measured using GNSS to track the x–y coordinates of the ASV. The advance measures the distance along the initial course up to a 90-degree heading change (t_q), and the tactical diameter measures the perpendicular distance across the initial course until a 180-degree heading change is achieved (t_h).

$$\text{Ad} = |x(t_q) - x(t_0)|$$

$$\text{TD} = |y(t_h) - y(t_0)|$$

For the test conducted at the specified speed, the tactical diameter should be less than five times the ship length (L_{ship}), and the advance should be less than four and a half ship lengths. The test should be repeated at different speeds to assess the MASS's turning ability across its entire operational speed range. According to the IMO standard, if Ad $< 4.5L$ and TD $< 5L$, the requirements for turning ability are met. However, stricter requirements may be set based on the MASS's intended purpose.

9.3.1.6 Zigzag tests

The Zigzag tests involve the vessel performing maneuvers that mimic the pattern of saw teeth by alternating the rudder position between predetermined positive and negative angles once the vessel's heading matches these angles. This test is designed to provide insights into the vessel's dynamic response and maneuvering capabilities, assessing how quickly and effectively it can change its course without significant overshoot.

The 10°–10° zigzag test involves setting the rudder to 10 degrees to either side, assessing the vessel's responsiveness and agility in dynamic environments. The key performance indicators for this scenario include the first and second overshoot angles and the time to reach the initial 10-degree heading angle change on both sides. Measurements of the heading angle ($\Psi(t)$) commence at the start of the maneuver.

$$\alpha_1 = \max\{\Psi(t)\} - 10^\circ, \quad \alpha_2 = \min\{\Psi(t)\} + 10^\circ$$

The distance traveled during the first overshoot should not exceed 2.5 ship lengths (L), calculated using GNSS coordinates measured at regular intervals until the first overshoot angle reaches its maximum.

$$l_{10} = \sum_{i=t_1+\Delta t}^{t_{\max}} \sqrt{(x(i) - x(i-1))^2 + (y(i) - y(i-1))^2} \leq 2.5L$$

The 20°–20° zigzag test increases the challenge by requiring the rudder to alternate at 20 degrees, testing the limits of the vessel's maneuvering system under more extreme conditions. This setup mirrors that of the 10°–10° zigzag, but the larger turning degrees place a greater demand on the vessel's maneuverability and the first overshoot angle becomes a more critical measure of the vessel's initial turning ability.

$$\alpha_1 = \max\{\Psi(t)\} - 20^\circ$$

This scenario provides a rigorous test of the MASS's ability to execute tight maneuvers, ensuring its suitability for operation in environments requiring high maneuverability and precise handling.

9.3.1.7 Pull-out test

The pull-out test assesses the dynamic stability of the autonomous ship and its ability to return to a straight course after a turning maneuver. This test, critical for evaluating how quickly the vessel stabilizes post-turn, can be conducted immediately following the turning test. Initially, the rudder is set to its maximum allowable angle from the end of the turning test, setting the stage for the pull-out test.

During this test, the rudder is quickly adjusted back to a neutral position and held there until the vessel achieves a steady turning rate. Ideally, a perfect performance shows a zero steady turning rate, indicating dynamic stability. If there is dynamic instability, the steady turning rate will show a residual value, where a higher residual suggests poorer performance.

The rate of heading change $\dot{\Psi}(t)$ is monitored from the start of the test (t_0) until a steady rate is achieved (t_s), using a gyro compass or GNSS. Dynamic stability is confirmed if $\dot{\Psi}(t_s) = 0$; instability is indicated by $\dot{\Psi}(t_s) \neq 0$, with increasing $|\dot{\Psi}(t_s)|$ signifying reduced performance.

9.3.1.8 Path following

Path following tests evaluate the MASS's ability to adhere to a designated route composed of multiple way-points, assessing the navigation system's effectiveness. These tests involve trajectories that the MASS navigates, represented as longitude and latitude coordinates.

The MASS's precision in passing way-points is crucial and measured using GNSS or visual aids like cameras. The primary metric is the deviation from the optimal path length.

The accuracy per way-point, W_{pi}, is determined by tracking the MASS's path using GNSS coordinates sampled at intervals δt, with the distance from $P(t)$ to W_{pi} minimized. The overall path accuracy is the average deviation from the way-points, which should be below a predetermined threshold a.

KPI 1:

$$\sum_{i=1}^{n} A_{ci} < a_1 \text{ for } n \text{ way-points, where } A_{ci} = \min_{\forall t} \sqrt{(x(t) - x_{Wpi})^2 + (y(t) - y_{Wpi})^2}$$

KPI 2:

$$d_t - OR < a_2, \text{ where } d_t = \sum_{j=t_0+\Delta t}^{end} \sqrt{(x(j) - x(j-1))^2 + (y(j) - y(j-1))^2}$$

9.4 Collision avoidance

Collision avoidance capability is critical for safe operations of autonomous ships, particularly in busy urban waterways and port environments. A robust collision avoidance system not only increases safety but also optimizes cost in terms of time and energy required to alter the vessel's path. According to international regulations, such as those outlined by the IMO's COLREGs, MASSs as well must adhere to specific guidelines that regulate how vessels should maneuver around each other to prevent collisions.

9.4.1 Avoid obstacle

The avoid obstacle test evaluates the autonomous surface vessel's collision avoidance capabilities by introducing both stationary and moving obstacles like other vessels into its path. This test is crucial for assessing the MASS's ability to detect and efficiently navigate around obstructions, maintaining safety without unnecessary delays or excessive energy use.

The primary KPIs are the registration distance (RD) – the distance at which the MASS first identifies an obstacle – and the efficiency of the evasive maneuvers, which compares the additional distance traveled during the maneuver to the planned route.

- **KPI 1: Registration Distance (RD)** must exceed the minimum safe distance (MD) to effectively avoid the obstacle.
- **KPI 2: Efficiency of the Evasive Maneuver** quantifies the total extra distance traveled during the maneuver compared to the initial route:

$$\sum_{i=t1+\Delta t}^{end} \sqrt{(x(i) - x(i-1))^2 + (y(i) - y(i-1))^2} - IR < a$$

  where *IR* is the length of the initial route and *end* marks the conclusion of the route.

9.4.2 *Overtaking*

In the overtaking scenario, an additional vessel moves at a lower speed along the same course as the MASS. The ship must overtake this vessel while maintaining a safe distance and optimal speed.

9.4.3 *Avoid crossing obstacles*

This test evaluates the MASS's response to a moving vessel on a crossing path, emphasizing timely and accurate decision-making to prevent collisions. The MASS must detect the crossing vessel's path, determine a collision course, and apply COLREG regulations to execute efficient evasive maneuvers.

Measurements track the additional distance traveled during maneuvers, the closest approach between vessels, and the distance at which the MASS first detects the crossing vessel. The successful application of COLREGs is crucial, assessing the MASS's ability to navigate safely with minimal impact on its route and energy use.

Key performance indicators include the following:

- **KPI 1: Registration Distance (RD)** must exceed the minimum safe distance (MD) for initiating evasive maneuvers.
- **KPI 2: Minimum Distance** ($\min\{d(t)\}$) between the MASS and the crossing vessel should surpass a set safety distance (SD).
- **KPI 3: Efficiency of the Evasive Maneuver** evaluates the extra distance traveled compared to the initial path:

$$\sum_{i=t1+\Delta t}^{end} \sqrt{(x(i) - x(i-1))^2 + (y(i) - y(i-1))^2} - IR < a$$

  where *IR* is the initial route's length and *end* is the maneuver's conclusion.
- **KPI 4: COLREG Compliance** checks if the MASS adheres to giving way if the crossing vessel is from the starboard side.

9.4.4 *Avoid head-to-head collision*

This scenario tests the MASS ability to handle situations where an oncoming vessel is heading directly toward it. The key challenge is to determine whether the vessels can pass each other without altering course or if evasive maneuvers are necessary. According to COLREGs, both vessels should ideally move to starboard to pass each other safely. This test assesses the MASS capability to correctly register and react to potential collision paths.

9.4.5 Avoid shallow water

The avoid shallow scenario examines the MASS depth perception and decision-making capabilities when encountering shallow waters or underwater obstacles. The risk of grounding is a critical concern, and the MASS must demonstrate the ability to effectively avoid such hazards. This scenario can be simulated by setting a course through shallow parts of a waterway or creating underwater obstacles that the MASS must detect and evade using historical data, GNSS, depth measurement tools or sonar.

KPIs for this scenario include the following:

- **KPI 1: Registration Distance (RD)** should be greater than the minimum distance (MD) required to initiate an evasive maneuver safely.

  $$RD > MD$$

- **KPI 2: Efficiency of Navigation** is measured by the additional distance traveled compared to the initial route:

  $$\sum_{i=t_1+\Delta t}^{t_{\mathrm{end}}} \sqrt{(x(i) - x(i-1))^2 + (y(i) - y(i-1))^2} - IR < a$$

  where IR is the initial route's length and t_{end} is the last time point of the route.

9.5 Vessel collaboration

The use of a fleet of autonomous vessels can significantly enhance efficiency and safety through coordinated operations. This section explores scenarios involving fleet formations, which are particularly beneficial in terms of energy consumption and control costs. The following vessels can utilize the wake of the leading vessel, reducing overall energy requirements, while simultaneously utilizing control outputs from the leader to maintain formation.

9.5.1 Keep constant distance

The "Keep Constant Distance" test examines an autonomous vessel's ability to maintain a set distance from another, a basic test for fleet formation capabilities. This test is essential for assessing the viability of operating MASS in a fleet, emphasizing safety and efficiency.

In this scenario, two MASS are used: one as the leader and the other as the follower. The objective is for the follower to autonomously adjust its speed and direction to keep a consistent distance from the leader as they navigate a predetermined course. This evaluates the fleet's navigational abilities and ensures that the maintained distance is safe to avoid collisions.

Distances between the MASS are continuously measured, with KPIs including the following:

- **KPI 1: Minimum Distance** ensures that the distance never falls below a safety threshold:

  $$\min\{d(t)\} > D - a_1$$

- **KPI 2: Maximum Distance** ensures that the distance does not exceed a set limit:

  $$\max\{d(t)\} < D + a_2$$

- **KPI 3: Average Distance** checks for consistent distance maintenance:

  $$D - a_3 < \frac{\sum_{t=1}^{n} d(t)}{n} < D + a_3$$

  with n being the number of measurement points.

9.5.2 *Form a fleet and keep formation*

This test evaluates the capability of at least three autonomous surface ships to form and maintain a specific formation. Initially, the MASSs are positioned within close proximity, with one designated as the leader. The leader sets a course while the others adjust their positions relative to the leader, optimizing fuel usage and computational load and leveraging aerodynamic or hydrodynamic benefits.

Key performance indicators include the average and extreme values of distances and angles between MASS, as well as the time taken to establish the formation. Precision in maintaining the formation is monitored using cameras, distance sensors, and GNSS.

Key performance indicators for maintaining the formation include the following:

- **KPI 1: Formation Time** measures the time to complete formation:

  $$t_f - t_0 < a_1$$

- **KPI 2–4: Distance Maintenance** ensures that distances between vessels stay within predefined limits:

  $$\min\{d_{ij}(t)\} > D_{ij} - a_2, \quad \max\{d_{ij}(t)\} < D_{ij} + a_3,$$

  $$\frac{1}{n}\sum_{t=1}^{n} d_{ij}(t) > D_{ij} - a_4 < D_{ij} + a_5$$

- **KPI 5-7: Angle Consistency** checks that angles relative to the leader are maintained:

  $$\min\{\theta_i(t)\} > \Theta_i - a_6, \quad \max\{\theta_i(t)\} < \Theta_i + a_7,$$

  $$\frac{1}{n}\sum_{t=1}^{n} \theta_i(t) > \Theta_i - a_8 < \Theta_i + a_9$$

Table 9.3 Overview of testing scenarios and their objectives

Scenario	Objective	Assessment aspects
Maneuvering	Test basic maneuvering and responsiveness.	Control, safety, navigability, reliability
Path following	Assess the MASS ability to follow a designated path using way-points.	Control, cost efficiency, navigability, reliability
Collision avoidance	Evaluate the MASS systems for avoiding collisions under various conditions.	Control, safety, cost efficiency, reliability
Vessel collaboration	Test the operation of MASS in a fleet, focusing on formation maintenance and dynamic repositioning.	Control, safety, navigability, energy efficiency

9.5.3 Change position in formation

This scenario involves two vessels within a fleet switching positions while maintaining formation integrity and safety, which can be critical in adapting to environmental changes or operational needs.

These collaborative scenarios are essential for testing the operational dynamics of MASS fleets, ensuring that they can perform complex maneuvers and maintain formations under various maritime conditions. The results from these tests validate the MASS readiness for coordinated and efficient fleet operations.

9.6 Overview of the scenarios

This subsection provides a comprehensive overview of all testing scenarios included in the guideline. Below is a summary table that outlines the objectives of each scenario and the corresponding assessment aspects. This summary helps to contextualize the reasons each scenario has been included in the testing method and how it contributes to the comprehensive assessment of the autonomous vessels.

9.6.1 Key performance indicators (KPIs)

Table 9.3 summarizes the KPIs for assessing the control and safety performance of MASSs across various test scenarios.

9.7 Conclusion

Throughout this chapter, we have explored various test scenarios designed to evaluate the performance and safety of MASS systems in diverse navigational contexts and introduced the RAS framework as a comprehensive approach for these assessments. These tests including maneuvers such as navigating in straight lines, mooring, reaching maximum speeds, stopping, turning, and zigzag paths are critical for ensuring that MASS operates efficiently and safely within their operational environments.

The implementation of collision avoidance tests underscores the importance of MASS' ability to interact with other vessels and stationary obstacles dynamically, ensuring compliance with established maritime regulations such as COLREGs. These tests not only demonstrate an autonomous ship responsiveness to immediate obstructions but also its ability to adhere to legal and safety standards that govern maritime operations.

Moreover, scenarios focusing on fleet formation and maintaining formation with multiple MASS emphasize the technological advances in cooperative navigation. These tests, summarized in Tables 9.4 and 9.5, evaluate the collective operational effectiveness and intervehicle communication within a fleet, marking a significant step toward integrated maritime traffic management systems.

9.7.1 Future recommendations

As the development of MASS continues to accelerate, several areas require further research and development:

- **Enhanced sensor integration:** Future work should focus on integrating more advanced sensor technologies to improve the detection accuracy and operational range of MASS. This includes the development of sophisticated LiDAR, sonar, and radar systems that can provide more detailed environmental feedback and better obstacle detection capabilities.
- **Robust control/AI-based algorithms:** There is a pressing need to develop more robust control and artificial intelligence algorithms that can better assist human decision-making in navigation and collision avoidance scenarios. These algorithms should be capable of handling high-stakes situations with minimal human oversight.
- **Intervehicle communication protocols:** Developing standardized protocols for intervehicle communication will be essential for the effective operation of MASS fleets. This will facilitate smoother coordination and operation across various types of maritime vessels.
- **Regulatory and legal frameworks:** As MASSs become more prevalent, updating regulatory frameworks to include provisions for autonomous operations will be crucial. This includes revising maritime laws to accommodate the unique aspects of unmanned vessels, ensuring that they can operate legally and safely alongside manned vessels.
- **Public acceptance and trust:** Efforts should also be directed toward building public trust and acceptance of autonomous surface vessels. This involves conducting extensive safety trials and transparently sharing results with the public and stakeholders in the maritime industry.

In conclusion, while autonomous ship systems hold substantial promise in revolutionizing maritime transport by enhancing operational efficiency and safety,

Table 9.4 Experimental scenarios and associated SI units

Category	Scenario	KPIs	Unit
Manoeuvring	Navigate in straight line	Average and maximum deviation from straight line	m
	Mooring test	Deviation of stopping location to set location	m
	Reach maximum speed	Maximum speed	m/s
	Stopping test	Track reach, Head reach	m
	Turning test	Tactical diameter, Advance	m
	10-10 zigzag	First overshoot angle, Second overshoot angle, Traveled distance during first overshoot	deg, m
	20-20 zigzag	First overshoot angle	deg
	Pull-out test	Residual turning rate	deg/s
Path following	Path following	Cumulative absolute deviation from the way-points, Deviation from the optimal path length	m
Collision avoidance	Avoid stationary obstacle	Registration range, Extra distance covered by performing maneuver	m
	Overtaking	Registration range, Minimum distance between the vessels, Extra distance covered by performing maneuver	m
	Avoid crossing obstacle	Registration range, Minimum distance between the vessels, Extra distance covered by performing maneuver, Correct implementation of COLREG	m, yes/no
	Avoid head-to-head collision	Registration range, Minimum distance between the vessels, Extra distance covered by performing maneuver	m
	Avoid shallow	Registration range, Extra distance covered by performing maneuver	m
Vessel collaboration	Keep constant distance	Minimum, maximum and average distance to leading MASS	m
	Form a fleet and keep formation	Minimum, maximum and average distance between MASS, Minimum, maximum and average angle between MASS	m, deg
	Change position in formation	Minimum distance between MASS, Time to execute change, Deviation from optimal route	m, s, m

efforts are still required to refine their technological capabilities and integrate them into existing maritime regulations and operational frameworks. A consistent set of performance indicators covering key aspects such as navigation accuracy, reliability, and environmental impact will be essential to monitor progress, guide

Table 9.5 Overview of test methods and KPIs

Scenario	KPIs	Measurement parameters	Measurement equipment	Assessment
Navigate in straight line	Maximum deviation; Average deviation	x–y coordinates over time	GPS; Camera	$\max\{d(t)\} < a1$; $\Sigma_n t = 0(d(t)/n < a2$
Mooring test	Deviation from location	x–y coordinates when moored	Distance sensor; GPS; Camera	$d < a$
Reach maximum speed	Maximum speed	x–y coordinates over time	Camera; GPS	$V_d - \max\{V(t)\} < a$
Stopping test	Track reach; Head reach	x–y coordinates over time	Camera; Distance sensor; GPS	$TR < 10L$; $\mathrm{HR} = \|x(t_1) - x(t_0)\|$
Turning test	Tactical diameter; Advance	x–y coordinates over time	Camera; GPS	$TD < 5L$; $Ad < 4.5L$
10°–10° zigzag	First overshoot angle; Second overshoot angle	Heading over time	Camera; GPS; Gyro	$\alpha_1 < 10^\circ$; $\alpha_2 > 10^\circ$
20°–20° zigzag	First overshoot angle	Heading over time	Camera; GPS; Gyro	$\alpha_1 < 10^\circ$
Pull-out test	Residual turning rate	Heading over time	Camera; GPS; Gyro	$\dot{\Psi}(t_s) = 0$
Path following	Cumulative deviation from waypoints; Deviation from path length	x–y coordinates over time	Camera; GPS; LiDAR	$\Sigma_n i = 1Aci/n < a1$; $dt - OR < a2$
Avoid stationary obstacle	Registration range; Extra maneuver distance	x–y coordinates over time	Camera; GPS; LiDAR	$RD > MD$; $Extradist. < a$

technological improvements, and align standards across the sector. Finally, continued collaboration among researchers, industry stakeholders, and regulatory bodies will also be important in realizing the full potential of autonomous maritime navigation.

References

[1] Abaei M.M. and Hekkenberg R.: "A method to assess the reliability of the machinery on autonomous ships." *19th Conf. on Computer Applications and Information Technology in the Maritime Industries*; 2020. pp 11–13.

[2] Aslam S., Michaelides M.P., and Herodotou, H.: Internet of ships: A survey on architectures, emerging applications, and challenges. *IEEE Internet of Things Journal*. 2020;**7**(10):9714–27.

[3] Benjamin M.R. and Curcio J.A.: "COLREGS-based navigation of autonomous marine vehicles." *2004 IEEE/OES Autonomous Underwater Vehicles (IEEE Cat. No.04CH37578)*; Sebasco, ME, 2004. Piscataway: IEEE. pp. 32–39.

[4] Chaal M., Bahootoroody A., Basnet S., Banda O.A.V., and Goerlandt F.: "Towards system-theoretic risk assessment for future ships: A framework for selecting Risk Control Options." *Ocean Engineering*. 2022; **259**:111797.

[5] Chaal M., Ren X., BahooToroody A., *et al.*: "Research on risk, safety, and reliability of autonomous ships: A bibliometric review." *Safety Science*. 2023;**167**:106256.

[6] Chae C.-J., Kim M., and Kim H.-J.: "A study on identification of development status of MASS technologies and directions of improvement." *Applied Sciences*. 2020;**10**(13):4564.

[7] Chang C.-H., Kontovas C., Yu Q., and Yang Z.: "Risk assessment of the operations of maritime autonomous surface ships." *Reliability Engineering & System Safety*. 2021;**207**:107324.

[8] Choi J., Park J., Jung J., Lee Y., and Choi H.-T.: "Development of an autonomous surface vehicle and performance evaluation of autonomous navigation technologies." *International Journal of Control, Automation and Systems*. 2020; **18**(3):535–45.

[9] Cockcroft A.N. and Lameijer J.N.F.: *Guide to the Collision Avoidance Rules*. Amsterdam: Elsevier; 2003.

[10] de Vos J., Hekkenberg R.G., and Banda O.A.V.: "The impact of autonomous ships on safety at sea–a statistical analysis." *Reliability Engineering & System Safety*. 2021;**210**:107558.

[11] Ellefsen A.L., Æsøy V., Ushakov S., and Zhang H.: "A comprehensive survey of prognostics and health management based on deep learning for autonomous ships." *IEEE Transactions on Reliability*. 2019;**68**(2):720–40.

[12] Garofano V., Hepworth M., and Shahin R.: Obstacle avoidance and trajectory optimization for an Autonomous Vessel Utilizing MILP Path Planning, Computer Vision based Perception and Feedback Control [Internet]. Institute of Marine Engineering, Science and Technology (IMarEST); 2022. Available from: http://dx.doi.org/10.24868/10713

[13] Goerlandt F.: "Maritime autonomous surface ships from a risk governance perspective: Interpretation and implications." *Safety Science*. 2020;**128**:104758.

[14] Hannaford E., Maes P., and Van Hassel E.: "Autonomous ships and the collision avoidance regulations: a licensed deck officer survey." *WMU Journal of Maritime Affairs*, 2022;**21**(2):233–66.

[15] Haseltalab A., Garofano V., and Afzal M.: The Collaborative Autonomous Shipping Experiment (CASE): Motivations, theory, infrastructure, and experimental challenges [Internet]. *Proceedings of the International Ship Control Systems Symposium (iSCSS)*. IMarEST; 2020. Available from: http://dx.doi.org/10.24868/issn.2631-8741.2020.014

[16] Hinostroza M.A., Xu H., and Soares C.G.: "Experimental and numerical simulations of zig-zag manoeuvres of a self-running ship model." In: Soares C.G. and Teixeira A.P. (eds), *Maritime Transportation and Harvesting of Sea Resources*. London: Taylor & Francis Group; 2018, pp. 563–70.

[17] Hirst H.: *COLREGS: Still Fit for Purpose?*; 2020.

[18] Hoem Å.S., Fjørtoft K.E., and Rødseth Ø.J.: *Addressing the Accidental Risks of Maritime Transportation: Could Autonomous Shipping Technology Improve the Statistics?* Poland: Gdynia Maritime University; 2019.

[19] Huang Y., Chen L., Chen P., Negenborn R.R., and Van Gelder, P.: "Ship collision avoidance methods: State-of-the-art." *Safety Science*. 2020;**121**:451–73.

[20] Organization International Maritime: *COLREG: Convention on the International Regulations for Preventing Collisions at Sea, 1972*. London: International Maritime Organization; 2003.

[21] Johansen T.A., Perez T., and Cristofaro A.: "Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment." *IEEE Transactions on Intelligent Transportation Systems*. 2016; **17**(12):3407–22.

[22] Kim M., Joung T.-H., Jeong B., and Park H.-S.: "Autonomous shipping and its impact on regulations, technologies, and industries." *Journal of International Maritime Safety, Environmental Affairs, and Shipping*. 2020;**4**(2):17–25.

[23] Madsen A.N., Aarset M.V., and Alsos O.A.: "Safe and efficient maneuvering of a Maritime Autonomous Surface Ship (MASS) during encounters at sea: A novel approach." *Maritime Transport Research*. 2022;**3**:100077.

[24] Mahacek P., Kitts C.A., and Mas I.: "Dynamic guarding of marine assets through cluster control of automated surface vessel fleets." *IEEE/ASME Transactions on Mechatronics*. 2012;**17**(1):65–75.

[25] International Maritime: *Standard for Ship Manoeuvrability, 1972*. London: International Maritime Organization; 2002.

[26] Negenborn R.R., Goerlandt F., Johansen T.A., *et al.*: "Autonomous ships are on the horizon: here's what we need to know." *Nature*. 2023;**615**(7950):30–3.

[27] International Maritime Organization: *COLREG: Convention on the International Regulations for Preventing Collisions at Sea, 1972*. London: International Maritime Organization; 2003.

[28] Orzechowski S.C.: “Regulatory scoping exercise for the future adoption of autonomous inland ships in Europe.” *Journal of Shipping and Trade*. 2024;**9**(1):2.

[29] Pang Y., Changyuan C., Garofano V., and Negenborn R.: *Assessment of Performance of Urban Freight Vessels.* Project co-funded by the INTERREG North Sea Region Programme 2014–2020 (ERDF). Publication in the Framework of AVATAR; 2023.

[30] Quigley M., Gerkey B., Conley K., *et al.* “Ros: an open-source robot operating system.” *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) Workshop on Open Source Robotics*, Kobe, Japan; 2009.

[31] Ramos M.A., Utne I.B., and Mosleh A.: “Collision avoidance on maritime autonomous surface ships: Operators’ tasks and human failure events.” *Safety Science*. 2019;**116**:33–44.

[32] Rødseth Ø.J. and Burmeister H.-C.: “Developments toward the unmanned ship” *Proceedings of International Symposium Information on Ships–ISIS*. Vol. 201, German Institute of Navigation (DGON), Hamburg; 2012. pp. 30–3.

[33] Statheros T., Howells G., and Maier K.M.: “Autonomous ship collision avoidance navigation concepts, technologies and techniques.” *Journal of Navigation*. 2008;**61**(1):129–42.

[34] Størkersen K.V.: “Safety management in remotely controlled vessel operations.” *Marine Policy*. 2021;**130**:104349.

[35] Tas J.: *Method for Testing and Assessing the Performance of Autonomous Surface Vessels.* Technical Report #2022.MME, 2022.

[36] Utne I.B., Rokseth B., Sørensen A.J., and Vinnem J.E.: “Towards supervisory risk control of autonomous ships.” *Reliability Engineering & System Safety*. 2020;**196**:106757.

[37] Åström K.J. and Murray R.M.: *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton: Princeton University Press; 2010.

[38] Öztürk Ü., Akdağ M., and Ayabakan T.: “A review of path planning algorithms in maritime autonomous surface ships: Navigation safety perspective.” *Ocean Engineering*. 2022;**251**:111010.