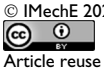
Special Issue: Autonomous Systems Safety, Reliability, and Security

Institution of MECHANICAL ENGINEERS

Journal of RISK AND RELIABILITY

Proc IMechE Part O:
J Risk and Reliability
1–23


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1748006X21051829
journals.sagepub.com/home/pio


A novel risk assessment process: Application to an autonomous inland waterways ship

Victor Bolbot[1], Gerasimos Theotokatos[1], Lars Andreas Wennersberg[2], Jerome Faivre[3], Dracos Vassalos[1], Evangelos Boulougouris[1], Ørnulf Jan Rødseth[2], Pål Andersen[4], Ann-Sofie Pauwelyn[5] and Antoon Van Coillie[6]

Abstract

Effectively addressing safety, security and cyber-security challenges is quintessential for progressing the development of next generation maritime autonomous shipping. This study aims at developing a novel hybrid, semi-structured process for the hazardous scenarios identification and ranking. This method integrates the operational and functional hazard identification approaches, whilst considering the safety, security and cybersecurity hazards. This method is applied to comprehensively assess the safety of an autonomous inland waterways ship at a preliminary design phase. The hazardous scenarios are identified and ranked by a number of experts participating in a series of sessions. The identified hazards risk is estimated considering the frequency and severity indices, whereas their uncertainty is estimated by employing the standard deviations in these two indices among the experts ranking results. Epistemic uncertainty is also considered during ranking. Risk control measures are proposed to de-risk the critical hazards. The results reveal that the most critical hazards from the safety, security and cybersecurity perspectives pertain to the situation awareness, remote control and propulsion functions. Based on the derived results, design enhancements along with high-level testing scenarios for the investigated autonomous ship are also proposed.

Keywords

Autonomous Inland Waterways ship, risk assessment, hybrid operational-functional process, safety, security and cyber-security, testing scenarios

Date received: 22 December 2020; accepted: 17 September 2021

Introduction

Paving the way towards the realisation of the Maritime Autonomous Surface Ships (MASS) requires innovative and bold initiatives. Such a collaborative initiative is the AUTOSHIP[1] project, which aims at converting two conventional ships (a Short Sea Shipping cargo ship and an Inland Waterways (IWW) barge) to autonomous ships and demonstrate the remotely controlled and autonomous ship operations at full scale conditions.

An important objective for the development and acceptance of MASS operations is to ensure their safety, cybersecurity and security. The safety challenges are attributed to the increased system complexity, as well as the involved interactions between the autonomous ships systems, subsystems and its environment.[2] Furthermore, cybersecurity needs to be addressed, as a successful cyber-attack could exploit vulnerabilities in the communication links and directly affect the integrity/availability of the data and control systems, leading to accidents.[2–5] A number of incidences with

[1]Maritime Safety Research Centre, Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, Glasgow, Scotland, UK
[2]SINTEF, Trondheim, Norway
[3]Bureau Veritas Marine & Offshore, Paris, France
[4]Kongsberg Maritime, Buskerud, Norway
[5]De Vlaamse Waterweg, Brussels, Belgium
[6]Zulu Associates, Kapellen, Belgium

Corresponding author:
Victor Bolbot, Maritime Safety Research Centre, Department of Naval Architecture, Ocean and Marine Engineering, University of Strathclyde, 100 Montrose Street, Glasgow, Scotland G1 1XQ, UK.
Emails: victor.bolbot@strath.ac.uk; vabolbot@gmail.com

unauthorised people gaining remote access to the control systems of conventional ships was recently reported in Wingrove.[6] Terrorists or pirates could potentially hijack an autonomous ship, taking over its control and subsequently attempt to collide with passenger/cruise ships and ports or demanding significant ransom.

Furthermore, except for a number of hazards related to the functional failures, external factors such as failures in another ship[7] or an emerging submarine[8] may also lead to incidents/accidents. Therefore, it is necessary to consider both the environmental context and the internal factors during the autonomous operations. In addition, hazards must be identified as early as possible in the design phase, so that appropriate design decisions are made. Nevertheless, the effectiveness of the hazard identification process is also an important requirement, as the involved partners (designers, owners, shipyards and classification societies) have limited resources. Therefore, it would be beneficial to standardise and automate/semi-automate the safety assessment process.[2] The lack of statistical data for the autonomous and unmanned ships also impedes the quantitative estimation of the risks and the associated uncertainty.[9] Another challenge is associated with the identification of relevant testing scenarios for autonomous ships. At the same time, the ship operating phases considerably affect the severity of the potential hazardous scenarios. For example, a ship blackout has practically no safety implications when the ship is anchored; however it might lead to collision, contact or grounding in the case of ship manoeuvring or sailing close to shore.[10]

Several approaches can be employed for the hazard identification at the initial stages of autonomous ships design. The first approach, which is typically employed by the classification societies[11,12] and regulatory authorities,[13] includes the implementation of Hazard Identification (HAZID) to identify and rank the hazardous scenarios as well as to verify the proposed design. This and similar approaches were also employed in various research studies.[14–17] The second approach employs systemic methods, such as System Theoretic Process Analysis (STPA)[18–28] and the third includes the adoption of standards, procedures and guidelines from other industries, for example, from ARP 4761[29] or ISO/PAS 21448.[30] Alternatively, approaches based on hazard lists and task analysis[31] as in ISO 10218[31] could be employed, but their application to autonomous ships has not been reported in the pertinent literature. Novel alternative methods can be also employed.

Whilst a number researchers consider the STPA a robust way to identify hazards on autonomous ships,[26] STPA has several limitations. STPA is based on the control structure and the identified list of hazards, without recommending a systematic way to identify these hazards. The processes described in the ARP 4761 standard focus on functional failures and therefore require separate analysis for hazards caused by external factors. Moreover, ARP 4761 does not consider security and cybersecurity. The ISO/PAS 21448 standard focuses on specific automated cars functions, not considering the complete car, and does not describe a specific methodology for the hazards identification. The methods presented in the ISO 10218 were developed for robotic applications.[31] Consequently, the application of a HAZID method (term frequently used interchangeably with the Preliminary Hazard Analysis[32]) to consider all ship functions for the identification of hazards at the initial design stages is considered as an effective approach. However, as reported in the pertinent literature, HAZID lacks structure[32] or focuses on the functional scenarios without sufficiently considering the contextual factors,[11] whereas safety is not thoroughly integrated with security and cybersecurity analyses. In addition, only a limited number of studies conducted risk assessments for the autonomous ship hazardous scenarios as reported above.

From the preceding analysis, it is concluded that for addressing the autonomous ships risk assessment at initial design stages, a comprehensive hazard identification and risk assessment process that considers the safety, security and cybersecurity, as well as the potential causes and consequences in various operating phases is required. Therefore, the aim of this study is to develop a structured process for the risk assessment of autonomous ships applicable at the initial design stages. This process is developed in such a way to allow for its semi-automation, thus rendering its implementation effective.

The novelty of the present study stems from: (a) the hybrid and semi-structured approach proposed for the hazard identification process taking into account both operational and functional hierarchical classification; (b) the application of the proposed novel hazard identification process for the case of an Inland Waterways ship, which reveals a considerable number of hazardous scenarios; (c) the proposal of a novel way to consider the uncertainty of the hazardous scenarios ranking; and (d) the comprehensive safety analysis integrated with the security and cybersecurity risk assessments (compared with pertinent studies that consider only a limited number of hazardous scenarios).

The remaining of this article is organised as follows. Section 2 describes the proposed methodology. In Section 3, the high-level information about the investigated ship is provided. In Section 4, the results of applying the novel HAZID process on an inland waterways ship are provided and discussed. Lastly, the main conclusions and findings of this study are reported.

Methodology description

Risk assessment process overview

The steps of the developed HAZID and risk assessment process are presented in the flowchart of Figure 1. The process follows the guidance for risk management according to International Standard Organisation (ISO) 31000[33] and it is also aligned to an extent with

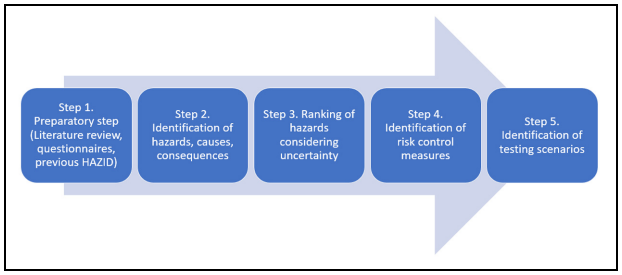


Figure 1. Hazard identification and risk assessment process overview.

the Bureau Veritas (BV)[11] guidance for autonomous ships risk assessment. However, it is modified to: (a) inherently integrate the security and cybersecurity risk assessment; (b) render the approach more systematic by using specific guidewords; and (c) include a novel way for analysing uncertainty. This process is proposed to be applied at the step 4.8 of the IMO[34] guidance for alternative systems approval.

First, the relevant safety information for the investigated system is gathered (step 1). In step 2, the relevant risks, their causes and consequences are identified. In step 3, the risks are analysed and ranked. In step 4, the risks are treated by relevant control measures. In the last step, testing scenarios for the autonomous ship verification are proposed.

Step 1: Preparatory step

The preparatory step aims to acquire and aggregate the required information about the ship as well as to develop a high-level supportive description of the ship systems. The main ship functions as well as the expected system preliminary design (ship and remote control centre), their input/output, responsibilities and interactions are identified. Moreover, the ship autonomy degree, operating phases and operating area are specified, whereas ambiguities with respect to the ship design are clarified. In addition, the ship functions, which will be automated, and their automation control and degree are specified.

Information on the investigated autonomous ship hazards is acquired by employing semi-structured interviews, where operators, authorities, system providers and original equipment manufacturers provided information on the expected hazards and risks. Previous studies on safety analyses with different structure are reviewed. Accident statistics for existing conventional ships are also investigated. Pertinent guidelines developed by a number of organisations such as BV,[11] United Kingdom chamber of shipping,[35] DNV GL[12] and Lloyds Register[36] are reviewed.

Step 2: Identification process for hazards, causes and consequences

The identification process for hazards, causes and consequences is implemented in a semi-structured way,

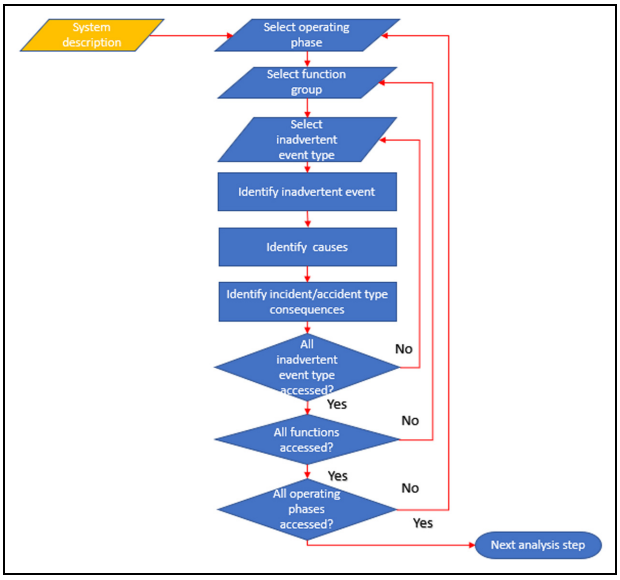


Figure 2. Flowchart of the developed hazard identification process taking into account a hybrid operational-functional approach.

according to the flowchart shown in Figure 2. To address the limitations discussed in the Introduction section, a hybrid approach that considers operational and functional hierarchical classifications is proposed in this study. The operational classification focuses on the different ship operational phases, whereas the functional classification focuses on the ship functions analysis. The process commences with a selected ship operating phase. Subsequently, a specific function group for this operating phase is selected for the analysis. For the combination of operating phase and function, a specific type of inadvertent events is considered either primarily related to safety or security or cybersecurity.

Herein, an integrated (rather than parallel) process is followed to allow for the simultaneous investigation of safety, security and cybersecurity issues, thus shortening the required time for the hazard identification and analysis. This was based on the argument that safety is a system property interdependent to cybersecurity and security.[37] Cybersecurity and security analyses were also included in this risk assessment process, so that causes to the various hazards are identified by describing a specific attack type. In cases where the cybersecurity/security breaches events have no direct impact on safety (although other consequences (financial/reputational) may be exhibited), they are considered as inadvertent events. In case where a safety related factor (e.g. failure) affects the security or cybersecurity inadvertent events, it is included as a cause to the relevant hazard as illustrated in Figure 3.

For each safety related event, accident/incident types (fire, collision, contact, grounding, explosion, machinery damage, foundering, personnel injuries)[38] are used to identify hazards. Herein, accident means any

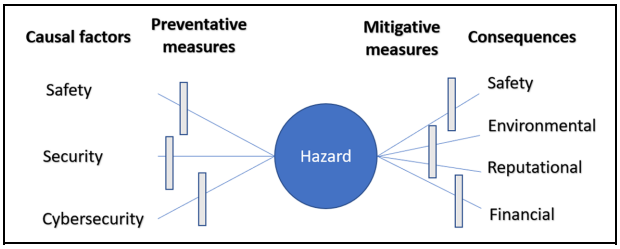


Figure 3. Dependencies between safety, security cybersecurity, hazards, consequences and mitigative barriers.

unintended (or intended) event involving fatality, injury, ship loss/damage, other property loss/damage or environmental damage.[39] Moreover, accident includes the events involving confidentiality, integrity and data availability loss as well as criminal activities, since safety, security and cybersecurity are jointly analysed. Herein, hazards mean 'the system states or the set of conditions that together with a worst-case set of environmental conditions will lead to an accident'.[40] Therefore, in the description of hazards, specific failures such a Closed Circuit Television cameras failure were avoided. Instead, generic conditions/states are used. The causes for the hazards though become very specific, as they depend on the system architecture description. The hazards are identified using guiding words as no output, wrong output (too much/too little), wrong input, untimely function output, conflicting output from function and asking 'what if' questions. To identify the hazards in specific operation, the functions in combination with the guidewords are being used by considering the potential accident types. The existing hazard lists are used for the verification and enhancement of generated hazard list.

For cybersecurity related events, in addition to the accident/incidents type, events as confidentiality breach and operation disruption are used. Furthermore, different threat groups (cyberterrorists, cybercriminals, hacktivists, generic hackers, competitors, states)[9] are considered and guidewords, such as steal (targetting at confidentiality), destroy, switch off (targetting at availability), get control over, transfer, manipulate, falsify (targetting at integrity) are used to identify relevant cybersecurity hazardous scenarios. In this way, confidentially loss, integrity loss or/and unavailability due to cyberattacks constitute either hazards or causes to hazards and are located either in the centre or at the left side of the Bow-Tie diagram shown in Figure 3.

For security issues, unauthorised access is used as the accident/incident type. Terrorists, hostile acts by insiders, organised crime, hooligans, competitors, states threat groups and guiding words as damage, hostage, hijack, seize of cargo, unauthorised access, smuggling weapons or drugs, using ship as weapon, attacks whilst on berth/sea, impede are used to identify the security hazardous scenarios.

For each safety, security, cybersecurity inadvertent event, potential causes are considered in terms of safety (system error, failure, human error, inadvertent environment conditions, missing input, management error), security (management failure, missing/faulty technical barrier, operating in dangerous area) and cybersecurity (potential cyberattacks types and vulnerabilities). For each of these scenarios inadvertent consequences are identified in terms of safety, damage to environment, reputational impact or financial impact. Finally, the format presented in Table 1 is developed and employed in this study for the identification of causes and consequences of each hazardous scenario. The initial hazards list is developed and reviewed by a number of experts during several HAZID workshops, where the hazards list is enriched and revised.

Step 3: Ranking of hazards considering uncertainty and risk

For ranking the identified hazardous scenarios, various risk tables can be used. The IMO[38] Formal Safety Assessment (FSA) risk matrix, which is widely employed in the maritime industry, was considered as basis for the purposes of this study. The frequency and severity are also ranked based on the respective categories provided in Tables 2 and 3, which were adapted by considering the guidelines from IMO,[38] BV,[11] DNV GL[41] and EMSA.[42]

The identified scenarios are initially ranked considering the investigated ship design as described in step 0, without taking into account risk control measures/options. The frequency is determined based on the causes and their likelihood of resulting into a hazard, whilst the severity is determined based on the expected consequences considering the classification of Table 3. Separate ranking is provided for the safety, security and cybersecurity causes of the hazards as well as the safety, environmental, reputational and financial consequences.

As autonomous ships constitute novel designs for which statistical data is not available, expert ranking was employed in line with the FSA guidelines.[39] The mean values for the frequency (FI) and severity (SI) rankings of each hazard are calculated according to the following equations:

$$FI^{j}_{Average} = \frac{\sum_{i}^{N} FI^{j}_{i}}{N_{FI}} \quad (1)$$

$$SI^{k}_{Average} = \frac{\sum_{i}^{N} SI^{k}_{i}}{N_{SI}} \quad (2)$$

where FI^{j}_{i} and SI^{k}_{i} are the rankings provided by the ith expert for the jth cause ($j = 1$ safety causal factor, $j = 2$ security, $j = 3$ cybersecurity) and the kth consequences ($k = 1$ safety consequences; $k = 2$ environmental

Table 1. Table describing hazards, causes and consequences.

No.	Operating phase	Fun. Id	Function	Incidents/hazards description			Causes			Consequences			
				Inadvertent event type	Accident type	Hazard	Safety related	Security related	Cybersecurity related	Safety	Environmental	Reputational	Financial
33	Transit	S11	Situation awareness (observation)	Safety	Collision	Other ship not activating its AIS/lighting system/ communication system	Other ships equipment failure AIS not available on other ships	Crew on other ships switching off its systems	AIS signal interference in the area	Collision with other ships	Leakage of stored bunker fuel	Negative coverage in the media	Damages to the own ship

Table 2. Frequency index (FI) rankings.[38]

Ranking (FI)	Frequency	Definition	Frequency (per ship−year)	Frequency (per ship−hour)
7	Frequent	Likely to occur once per month on one ship	10 (5−50)	$1.14\ 10^{-3}$
5	Reasonably probable	Likely to occur once per year in a fleet of 10 ships, that is, likely to occur a few times during the ship's life	10^{-1} ($5\ 10^{-2}$−$5\ 10^{-1}$)	$1.14\ 10^{-5}$
3	Remote	Likely to occur once per year in a fleet of 1000 ships, that is, likely to occur in the total life of several similar ships	10^{-3} ($5\ 10^{-4}$−$5\ 10^{-3}$)	$1.14\ 10^{-7}$
1	Extremely remote	Likely to occur once in the lifetime (20 years) of a world fleet of 5000 ships.	10^{-5} (0−$5\ 10^{-5}$)	$1.14\ 10^{-9}$

Table 3. Rankings for severity index (SI) (consequences) based on FSA,[38] BV,[11] DNV GL RP A-203 guidelines[41] and EMSA[42] report.

Ranking (SI)	Severity	Safety	Environmental			Financial	Reputation
		Effects on human safety	Oil spillage definition	Air pollution	Other for example, for ballast water treatment failures or collision with sea animals	Effect from ship operation disruption/ court costs/insurance costs/fines/effect on ship repair (treated as all together)	Effect on company reputation
5	Catastrophic	Multiple fatalities (1−10 and more)	Oil spill size between 100 and 1000 t or more than 1000 t	Major air pollution with long-term environmental consequences	Impact such as persistent reduction in ecosystem function or significant disruption of a sensitive species	$80,000,000 (⩾$25,000,000) Total loss	Extensive negative attention in international media/ industry
4	Severe	Single fatality or multiple severe injuries. Full recovery with extensive medical treatment	Oil spill size between 10 and 100 t	Air pollution resulting in air evacuation	Impact such as significant widespread and persistent changes in habitat, species, or environment media	$8,000,000 ($2,500,000 −$25,000,000) Severe damage	National impact and public concern; Mobilisation of action groups
3	Significant	One or more injuries, not severe. Full recovery with medical treatment	Oil spill size between 1 and 10 t	Limited environmental impact due to air pollution involving reporting to authorities	Impact such as localised but irreversible habitat loss or widespread, long-term effects on habitat, species or environmental media	$800,000 ($250,000 −$2,500,000) Non-severe ship damage	Considerable impact; regional public/slight national media attention
2	Minor	One or more first-aid injury. Treatment is minimal or not necessary.	Oil spill size ⩽ 1 t	Limited to no air pollution	Impact such as localised, long-term degradation of sensitive habitat or widespread short-term impacts to habitat, species or environmental media	$80,000 ($25,000−$250,000) Local equipment damage	Limited impact; local public concern may include media
1	Negligible	Minor first-aid injury to a single person in the workforce. Treatment is minimal or not necessary.	Non-significant spill	Minor environmental impact	Impact such as localised or short-term effects on habitat, species and environmental media	$8000 (⩽$25,000) No damage	Slight impact; local public awareness, but no public concern

Table 4. Uncertainty level (UL) ranking.[34,45]

Uncertainty level (UL)	Assumptions	Novelty of involved technology	ΔRI
Low (1)	Reasonable assumptions	Use of proven technology (BV technology rating 0/IMO degree of novelty 1)	According to equation (6)
Medium (2)	Reasonable assumptions made, although simplifying the phenomena	Limited field history of technology or use of proven technology in novel environment (BV technology rating 1/IMO degree of novelty 2)	According to equation (6)
High (3)	Poor justifications for the assumptions made	Novel technology use or use of technology with limited history in novel environment (BV technology rating 2–3/IMO degree of novelty 3–4)	$\Delta RI = 1$

consequences; $k = 3$ reputational consequences; $k = 4$ financial consequences), respectively, whereas $N_{SI} = N_{FI}$ denotes the total number of experts who provided the respective rankings for each hazard.

The agreement among experts for the causes of hazard frequency and consequences of each hazard ($FI^j_{Average}$ and $SI^j_{Average}$) is quantified according to the equations (3) and (4), which represent the deviations from the average value (standard error of the mean).[43] The estimation of agreement is implemented to achieve better understanding of the risk uncertainty, which is a crucial part of the risk as described by Kaplan and Garrick[44] This formula is widely used during physical experiments for the estimation of error in the mean value of estimated physical parameter. The working assumption behind this formula is that all the experts' rankings are considered as 'test measurements' of the FI and SI and follow the normal distribution. So, the more experts are involved the more accurate will be the results in similar manner with experimental physics, where the more measurements are undertaken, the smaller the error in the mean value is. Also, in the analysis all the experts are treated as equally important, without assigning any specific weight. Such a consideration has been incorporated, since a novel technology is studied in the analysis and it would be challenging to consider that somebody from the involved participants, has more knowledge than the rest of the experts in the group. The advantage of these formulae is their relevant simplicity. To the best of authors knowledges this formula hasn't been used for treatment of experts' rankings in maritime risk assessment.

$$\Delta FI^j = \frac{\sigma_{FI^j}}{\sqrt{N_{FI}}} \tag{3}$$

$$\Delta SI^k = \frac{\sigma_{SI^k}}{\sqrt{N_{SI}}} \tag{4}$$

The σ_{FI} is standard deviation in the analysis.

The Risk index $RI_{Average}$ is estimated based on the traditional risk definition (considering that the indices correspond to the logarithms of the risk, frequency and severity, respectively), according to the following equation, based on the FSA guidance[38]:

$$RI_{Average} = \max(FI^j_{Average}) + \max(SI^k_{Average}) \tag{5}$$

The RI standard error is subsequently estimated by using the average errors of FI and SI, according to the following equation:

$$\Delta RI = mean(\text{mean}(\Delta FI^j), \text{mean}(\Delta SI)) \tag{6}$$

For ranking the uncertainty in this study, except for the agreement among the experts described previously, the uncertainty level (UL) of the involved novel technology is evaluated according to the categories shown in Table 4, which are based on IMO[34] and BV[45] guidelines for assessing novel technologies. A set of criteria for ranking the employed assumptions is taken from Flage and Aven[46] as well as Goertland and Reniers.[47] The ranking of the assumptions and technology novelty is conducted by the experts in parallel to the ranking of hazards and inadvertent events for each hazard. If the epistemic uncertainty is high (average UL higher than 2), then the ΔRI is taken at least 1 or greater if the agreement among experts is already small.

The scenarios exhibiting the highest sum of RI and ΔRI (equation (7)) are classified as critical scenarios. This extra precautionary measure (by adding ΔRI) is introduced to avoid any negative overwhelming consequences due to improper ranking and uncertainty in rankings. This treatment is also in line with ISO 31000, which defines the risk definition as 'the effect of uncertainty on objectives'.[33]

$$RIU = RI_{Average} + \Delta RI \tag{7}$$

Step 4: Identification of risk control measures

This step focuses on identification of relevant risk control measures. For all the hazards, mitigation measures

(fail-safe procedures or minimum risk conditions, depending on terminology) are identified.

Both preventative and mitigative control measures are specified only for the more risky scenarios identified in step 2. The risk can be reduced by considering the following categories of risk control measures,[33,11] which are classified as design, engineering, operational or financial: (a) designing out risk; (b) using safety devices; (c) applying fault tolerance techniques; (d) developing operational procedures and training; (e) avoiding risk; and (f) sharing risk.

The initial control measures are proposed by the process facilitators using available information, such as BV[11] guidelines, and subsequently reviewed by the participants of the HAZID workshops. These measures are considered as preliminary for the investigated design. For the critical hazards, it is proposed that more detailed safety methods are used to identify the safety/security/cybersecurity issues, which would also support the critical hazards analysis in more detail.

The control measures are also ranked based on their cost effectiveness (the cheapest solutions will be preferred), risk reduction and maturity (the most commercial solutions will be used). The maturity of the control measures is ranked according to third column of Table 4 in terms of the technology novelty. The ranking of the costs is implemented according to Table 3.

Step 5: Identification of testing scenarios

In the last step of the risk assessment process, the relevant hazards along with the preventative and mitigative risk control measures are used to derive the initial set of testing scenarios. This is in line with the BV[11] guidance for autonomous ships, but also in line with the Vee design process[29] and the current design process for cyber-physical systems, which requires the identification of the testing scenarios.[48–50] The mitigative control measures are used to derive the testing scenarios, which are necessary for the ship to employ fail-safe (minimum risk) conditions. The preventative scenarios are used to demonstrate the critical functionalities of the initial system. These testing scenarios are generic and need to be refined using more advanced safety methods, so that specific test scenarios are proposed. However, this is within the scope of a subsequent design phase.

Case study description

This study employs the case study of an existing Inland Water Ways (IWW) barge, considering its theoretical next-generation autonomous design including the ship and its systems as well as the Remote Control Centre (RCC). The description of this integrated autonomous system is carried out based on information acquired from the pertinent literature[3,51–56] and the AUTOSHIP project deliverables,[57] as well as feedback received from AUTOSHIP partners. The main particulars of the existing IWW ship, which will be used as demonstrator in the AUTOSHIP project, are provided in Table 5, whereas a picture of this ship is illustrated in Figure 4. It must be noted that the demonstrator of the AUTOSHIP project and the case study autonomous system (ship and its RCC), albeit share some similarities, differentiate in the considered installed systems/sub-systems and autonomy degrees.

Figure 4. Zulu 4 IWW barge.[58]

Table 5. IWW main particulars.

Property	Value/reference (unit)
Length	50 m
Breadth	6.6 m
Sailing speed	17 km/h
Draft – fully loaded	1.9 m
Carrying capacity design	300 t

The investigated case study considers an Autonomy Degree Three (or above) according to IMO[59] guidelines. This pertains to: 'Remotely controlled ship without seafarers on board, whereas the ship is controlled and operated from another location'. Furthermore, the investigated case study can be classified at level 3 according to CCNR,[60] which corresponds to constrained autonomous crewless ship operation.

Conventional IWW barges are primarily operated at inland waterways within Belgium and the Netherlands. Future operation is considered in all waterways of member states of the European Union, as well as Switzerland, UK and Norway.

Based on the information from ship owner, the cargo and the ship capital values were estimated to USD150k and USD1.5M, respectively. It must be noted that these numbers are only rough approximations and they do not correspond to the respective accurate values; however, they are used to indicate the scale of these costs.

The considered power and propulsion plant layout for the investigated IWW use-case ship is illustrated in Figure 5. The existing ship has one aft thruster that is driven by a four-stroke diesel engine connected through a gear box. The ship has one main switchboard connected to one diesel-generator set, and one emergency genset connected to the main switchboard via the emergency switchboard. Power from the main switchboard to the bow thruster is controlled through the variable

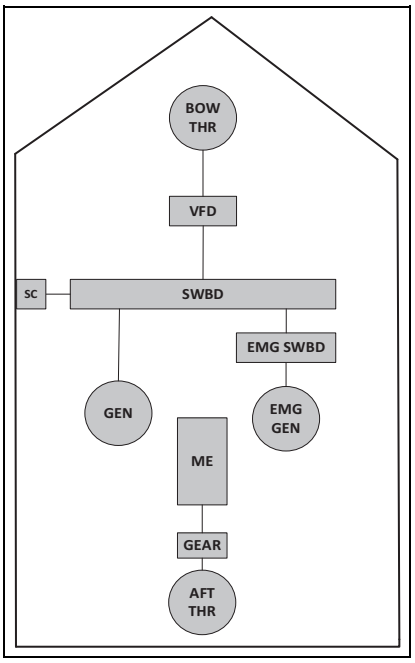


Figure 5. Power and propulsion setup of the IWW use-case ship.

frequency drive, whereas the shore electrical connection line is also connected to the main switchboard. Similar propulsion was considered herein for the use case, as the initial ship propulsion system is used as reference point. An overview of systems, sub-systems as well as their interconnections for the investigated case system (ship and RCC) is provided in Figure 6.

This study considers the following operating phases: (a) planning of the mission, (b) activities in ports, (c) arrival/departure from ports, (d) transit (including transit through locks), (e) emergency, (f) maintenance and repair (g) any (all) operating phases where the hazards are independent from the operating phase.

The analysis boundaries and stakeholders' groups are provided in Figure 7. The focus will be put on the ship and remote control centre related functions. The function groups were classified by employing the categorisation and breakdown developed in AUTOSHIP.[61] All the ship functions at Sxy level (xy is used for numbering the functions; x refers to group, y to subgroup) were considered in this analysis.

Several experts from various AUTOSHIP related organisations participated in two consecutive online workshops to carry out the hazard identification and ranking (first workshop), as well as the ranking of critical hazards considering the risk control measures (second workshop). The first workshop lasted 8 h, whilst the second lasted for 3 h. The details of the workshops' participants are provided in Table 6. We have included in the analysis the operator, who is knowledgeable of both the navigational and mechanical issues. It should be noted that the manned operation of the IWW involves only the captain (one person). There is no such rank as chief engineer, authorities, safety engineers and Original Equipment Manufacturers, cybersecurity experts. During each workshop, thorough discussion of the hazards and potential risk control measures took place. Prior to the workshops, information about the investigated case study ship and hazard identification process was distributed to the participants.

Results

Step 2: Identification hazards, causes and consequences

For the investigated IWW autonomous ship case study, in total 89 hazards were identified using the developed HAZID process as presented in Figure 8. During the hazards review with the involved experts, six of them were eliminated as not relevant or as duplicated. Some of the hazards are common between the transit and arrival operating phase. These were not eliminated as the consequences are different for each case. The complete list of hazards is available in Appendix B.

The typical causes related to the function groups are provided in Table 7. As it can be observed, several of the causes are repeated in different Sxy functions (xy is used for numbering the functions; x refers to group, y to subgroup). This was anticipated as several components are identical in the system layouts servicing different functions. However, this also indicates that some parts of the HAZID process could be semi-automated.

The list of the main consequences for the different operating phases are provided in Table 8. This list is smaller than the hazard list as it was anticipated that hazards ultimately lead to a limited number of accidents.

Step 3: Ranking of hazards considering uncertainty

Figure 9 presents the distribution of the evaluated risk index including the uncertainty level (RIU). It is observed that the vast majority of the identified scenarios have value of RIU below 7. Only seven scenarios exhibit a RIU above 8. The estimated high value of RIU is attributed to the dispersion exhibited among the experts' rankings as well the use of novel technology in specific scenarios.

As it can be observed from Figure 10 presenting the SI distribution, very few scenarios were considered to have high severity index (more than 4 on average involving fatality). These were primarily the scenarios associated with the ship navigation function (inappropriate situation awareness resulting in objects not being detected and collision with other ships). Such hazardous scenarios can lead to loss of life when other manned ships are involved. Explosion was also assigned high severity, as it can lead to significant damages to the ship as well as people and infrastructure in its surrounding. In this respect, the scenarios do not

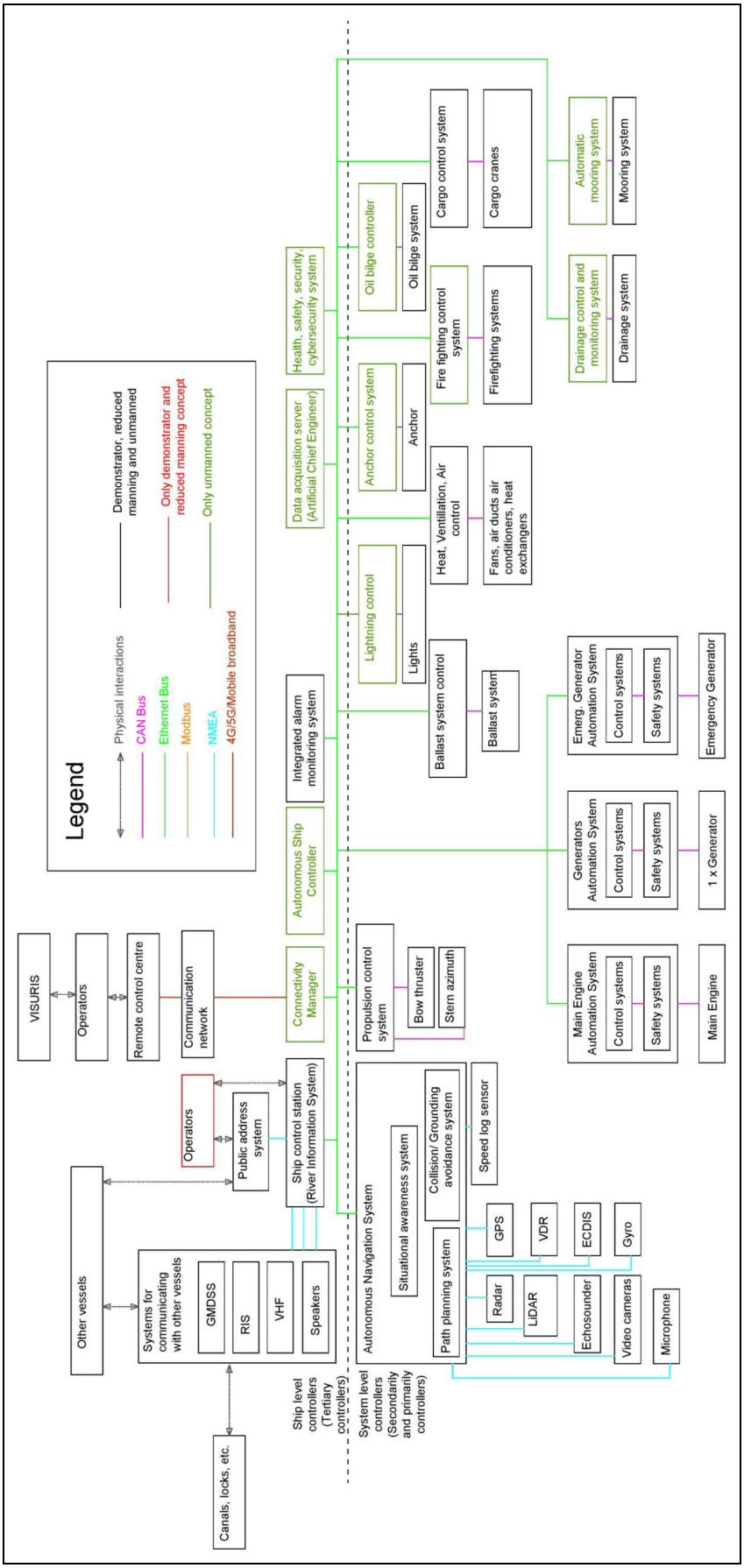


Figure 6. Ship systems and the communication network.

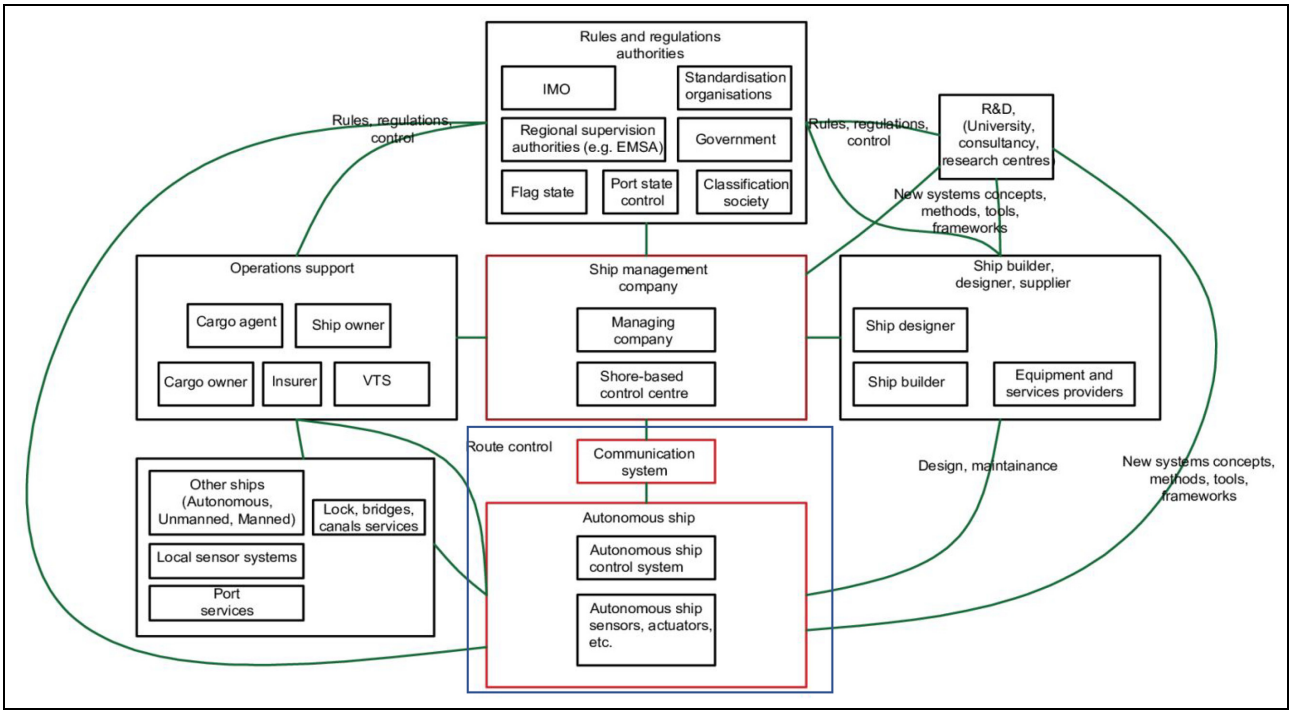


Figure 7. Autonomous vessel's system safety control structure (based on Wróbel et al.[18] and Puisa et al.[62]) and scope description (with red are highlighted the investigated parts in this study).

Table 6. Information about the workshops participants.

Participant no.	Expertise	Educational level	Experience
First workshop			
1	Academic in safety/security	PhD degree	> 20 years
2	Academic in safety/security	PhD degree	10–14 years
3	Safety engineer	MSc degree	> 20 years
4	Safety engineer	MSc degree	15–19 years
5	Original equipment manufacturer	MSc degree	> 20 years
6	Original equipment manufacturer	MSc degree	> 20 years
7	Original equipment manufacturer	MSc degree	> 20 years
8	Public authorities in safety	MSc degree	< 5 years
9	Academic	MSc degree	> 20 years
10	Public authorities in safety/security/cybersecurity	MSc degree	10–14 years
11	Safety engineer	PhD degree	> 20 years
12	Academic in safety/security/cybersecurity	PhD degree	< 5 years
Second workshop			
1	Academic in safety/security/cybersecurity	PhD degree	< 5 years
2	Academic in safety/security	PhD degree	10–14 years
3	Operator	MSc	> 20 years
4	Safety engineer	MSc	> 20 years
5	Original equipment manufacturer	MSc	> 20 years
6	Safety engineer	MSc	> 20 years
7	Authorities in safety	MSc	< 5 years
8	Authorities in safety/security/cybersecurity	MSc	10–14 years

differentiate from the respective ones for conventional ships, but the causes to these hazards are different.

The highest severity was estimated for the safety and reputational consequences involving risks to third parties or negative public coverage of the accidents as depicted in Figure 10. Instead, the consequences related to damages to ship/infrastructure or the operation disruption or environmental damage were generally ranked as less severe in the majority of cases. This can be attributed to the fact that the investigated ship is

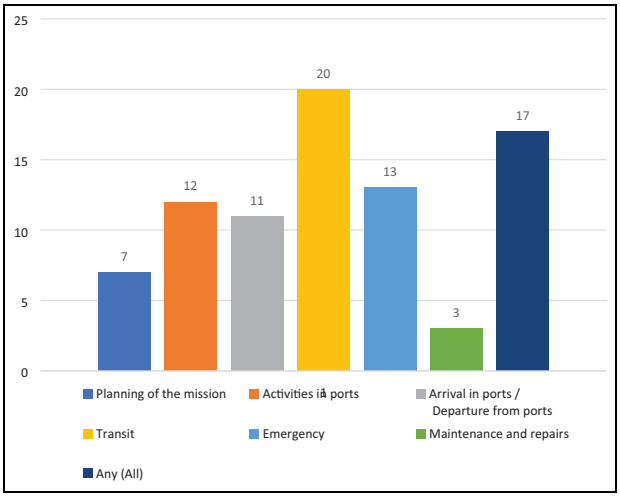


Figure 8. Hazards distribution per operating mode.

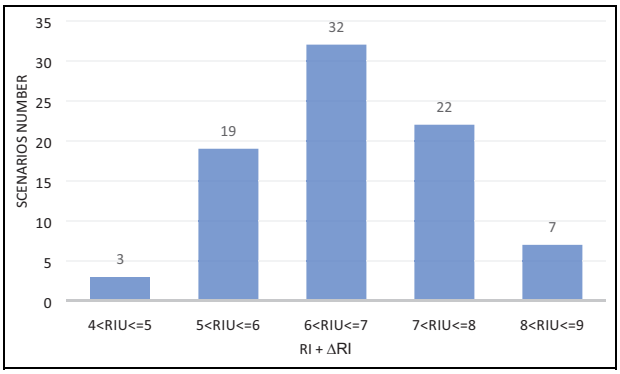


Figure 9. Distribution of risk index considering uncertainty (RIU = RI + UL).

small (hence the financial consequences of its loss are limited). It is noted that for the investigated ship, the possibility of fuel or lubricants leakage due to hull penetration is low, as cofferdams are used. In the unlikely event of a complete ship loss, fuel and lubricating oil can only leak through the tanks vents and sounding pipes, which can be controlled using appropriate isolation valves.

It should be noted that not all the scenarios were associated with all the consequences types (safety, environmental and other). Therefore, the number of scenarios for different consequences types (e.g. environmental) does not necessary sum up to 83. The reputational related consequences for the unmanned ships were found to be here more important compared with conventional manned ships, due to the novel technology use. For the unmanned ships that do not carry crew and passengers, the safety related risks are related to accidents involving other manned ship and port facilities.

The distribution of FI is provided in Figure 11. It is observed that the safety related causes mostly contribute to the risk in the investigated scenarios, as 27 of

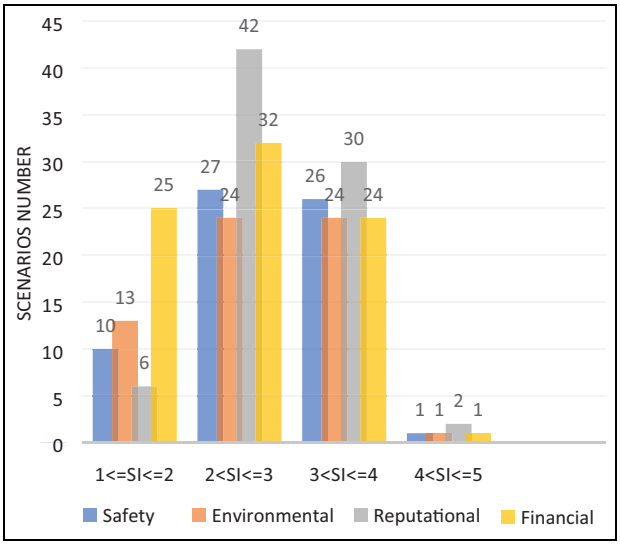


Figure 10. Distribution of severity index (SI) rankings.

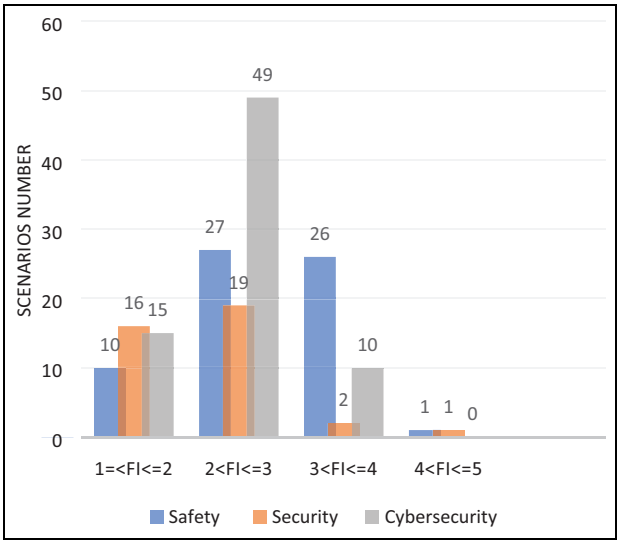


Figure 11. Distribution of frequency index (FI) rankings.

them have FI greater than 3 (characterised as remote according to Table 2). Only in 10 scenarios causes related to cybersecurity exhibit FI greater than 3, whereas three scenarios with causes related to security have FI greater than 3. The security, cybersecurity and safety related causes were associated with 38, 74 and 65 scenarios, respectively. It should be noted though that the number of cyber security experts participating in this study is limited, which might have influenced the results.

The highest FI values were associated with: (a) the thrust system failures; (b) difficulty in passing through locks; and (c) collisions with submerged large objects in canals, such as cars. The first is attributed to the fact that only one propeller was considered in the investigated IWW ship design. The presence of submerged objects, for example, cars or bikes at the bottom of

Table 7. Typical causes.

Code	Name of function group	IWW use typical causes		
		Safety related causes	Security related causes	Cybersecurity related causes
S11	Situation awareness (observation)	Inadequate training of algorithms Inadequate coverage of scenarios Failures in the equipment Environmental factors affecting the sensors performance (rain, rodents, organic waste, etc.) Damages to equipment due to interactions with port facilities Inadequate maintenance Fire (accident used as a cause) Flooding (accident used as a cause)	Acts of vandalism on the ship equipment	Cyberattacks on the situation awareness equipment Erroneous data coming from other ships
S12	Ship control (manoeuvring)	Failures in navigation system equipment Fire Flooding Inadvertent environmental factors (waves, currents, winds Water level, squat effects) Inadequate control from remote control centre	Unauthorised access to the ship control room	Cyberattacks
S13	Voyage management (navigation)	Errors in signals used for navigation	Unauthorised people in RCC or ship	Cyber-attacks
S14	Nautical communication	Failures in involved systems Environmental factors affecting the performance		Denial of service attack
S22	Mooring and anchoring	Failure in mooring equipment Failures due to interactions with other ships Fire (accident used as a cause) Flooding (accident used as a cause)		
S33	Hull integrity and strength	Lack/inadequate design of hull stress monitoring system, inadequate loading Inadequate inspection of ship structure Drain system failure		
S34	Stability and trim	Wrong loading Some additional loads Ballasting system failure		
S41	Power generation and operation	Failures Errors in design	Vandalism on the equipment	Cyberattacks
S42	Electrical systems	Arcs Short circuits Components failures Fire (accident used as a cause) Flooding (accident used as a cause)		
S43	Steering, propulsion and thrust	Components failures Fire (accident used as a cause) Flooding (accident used as a cause) Objects in propeller Sensors failure	Vandalism on the equipment	Cyberattacks
S45	System monitoring	Failure in system design Sensor failures CCTV failure		
S51	Security	Systems failures	Drugs/weapons trafficker accessing system	
S52	Cyber-security	Failure in management Software failures Lack of protection		
S53	Ship safety	Failure in management at RCC Software failures		
S54	Emergency management	Failures in firefighting system Failures in resuscitating equipment		Cyberattacks

Table 8. Main consequences for various operating modes.

a/a	Operating phase	Safety	Environment	Reputation	Financial
1	Mission planning	Collision with other ships leading to human injuries/fatalities	Fuel and lubricants leakages following a hull breach	Depending on the other consequences	Damages to ship infrastructure/ ship
2	Activities in ports	Involved personnel (locks/loading/ unloading) injury/ hostage			Ship unavailability
3	Arrival in ports/departure from ports	Contact with pier threating the lives of people on the peer Collisions with ships in vicinity leading to human injuries/ fatalities	Fuel and lubricants leakages following a hull breach		Ship/equipment damages/pier damage
4	Transit	Collision with passenger ship leading to human injuries/ fatalities			Foundering and shipwreck removal costs Ship/equipment damages Damages to other ships Damages to bridges locks
5	Emergency	Collisions leading to human injuries/ fatalities			Damages to ship and other ships
6	Maintenance and repair	Collision/contact groundings leading to human injuries/ fatalities			System damages
7	All phases	All the above	All the above		Ship damages/ fire

canals, is one of the inherent hazards for the IWW ships, that needs to be tackled by both manned and unmanned ships. The challenges associated with passing through locks are unique for unmanned ships due to the novel technology use, as they are related to the development of relevant technology.

As it can be observed from the Table 9, the majority of the high risk scenarios considering uncertainty are related to the ship navigation (directly or indirectly). The scenarios with high RIU that are directly linked to the navigation are: thruster loss, situation awareness system (used on the ship) failures, failure in communication systems and other ship switching off its AIS/ communication system or without AIS system such as kayaks. These scenarios also appear in manned ships (except the scenario H46), yet they can be attributed to completely different causes on the unmanned ship (mainly causes associated with the novel technology). The collision with bridges is considered as highly risky due to the significant infrastructure repair costs, which is similar to the manned ships cases.

It should be noted that inadvertent events related to confidentiality loss or stealing ship/cargo were not assessed to bear high risk (their RIU was found less than 7). This does not mean that the cybersecurity and security hazards are not important; still they need to be addressed sufficiently.

It is anticipated that automation on small ships can proceed faster due to reduced severity compared to large ships. Still, the risks to the navigation function are critical and need to be addressed effectively, as they may lead to significant human loss, for example, collision between unmanned ship and a passenger ship.

Step 4: Identification of risk control measures

The summary of selected risk control measures for the identified hazardous scenarios is provided in Table 10. Increased redundancy, such as redundancy in communication links, in the situation awareness system and propulsion systems can be used to prevent hazards from occurring. Some of these proposed measures are also applicable to the conventional ships. The financial consequences to a large extent are controlled by the insurance companies. The reputational consequences depend on the safety and environmental consequences, which implies that as long the other consequences are

Table 9. Most risky hazardous scenarios.

Operating phase	Id	Accident type	Hazard/(*guideword*)	Max (FI)	Max (SI)	Max (RI)	Δ***RI***	Max (RI) + Δ***RI***
				I/N	I/N	I/N	I/N	I/N
Transit	H32	Collision/ grounding/ contact	The surrounding situation is not properly (wrongly or not) determined/(*wrong/no output*) For example, objects are not detected and recognised (small objects, navigation marks, ships, ship lights, floating objects, depth, recreation crafts, people); weather is not properly conceived	3.9/3.8	4.1/2.8	8.0/6.6	1.0/1.0	9.0/7.6
Arrival in ports/departure from ports	H13	Collision/ grounding/ contact	Control loss over propulsion/thrust function/ (*wrong/no output*)	4.4/3	3.9/2.5	8.3/5.5	0.5/0.4	8.8/5.9
Transit	H33	Collision	Other ship not activating its AIS/lighting system/ communication system/ (*wrong input*)	4/3.6	3.9/2.8	7.9/6.4	0.6/1.0	8.5/7.4
Transit	H37	Collision	Ship on collision track with other ships in canal/(*wrong output*)	3.4/3.3	4.1/2.8	7.5/6.1	1.0/1.0	8.5/7.1
Transit	H46	Collision / Grounding/ Contact	Ship losing communication with the remote control centre/(*no output*)	3.7/3.0	3.6/2.6	7.3/5.6	1.0/1.0	8.3/6.6
Transit	H43	Collision	Loss of communication with other ships/(*no output*)	3.8/3.3	3.3/2.4	7.1/5.7	1.0/0.2	8.1/5.9
Transit	H36	Contact	Ship on collision track with bridge/lock/(*wrong output*)	3.3/3.1	3.7/2.4	7.0/5.5	0.8/0.4	8.0/5.9

I: initial ranking; N: ranking after control measures considered.

confined, the reputational consequences will be also acceptable.

The updated risk index for the critical hazardous scenarios after the implementation of the risk control measures is also provided in Table 9. The RI and the uncertainty are in general reduced for the critical hazards/inadvertent events. For all hazards, the SI can be reduced through the provision of relevant mitigation risk control measures. Although, the FI reduction was also anticipated, the results did not demonstrate a strong reduction trend. This is attributed to the lack of sufficient data for autonomous ships and the experiential nature of the ranking process as well as the fact that the two workshops were attended by a slightly different number of experts. Nonetheless, the estimated average FI variation lays within the uncertainty region, which was found around 1 for some of the scenarios (due to use of novel technology). This is unavoidable when experts ranking is used, especially in this case when pertinent data is not available.

The exhibited FI slight reduction is also attributed to the specific scenarios. For hazard H32 (identified as critical), the situational awareness is a function involving novel systems and sensors. Hence, the proposed measures were also considered during the initial hazard ranking, as they constitute part of the investigated autonomous system design. Since this hazard is linked to novel technologies, it is anticipated that it will be associated with high frequency and uncertainty. As the proposed risk preventative measures are of low maturity, the risk mitigation measures are required to for confine this risk. Similar comments apply to the hazards H33, H43, H36 and H37, which are linked to novel technologies. For H13 one level reduction in FI was observed (from 4.4 to 3). This indicates that incorporation of the redundant thruster system is anticipated to reduce the frequency of propulsion loss significantly in the unmanned ship compared to manned ships which use single thrusters.

It was found that the collision avoidance and situation awareness functions control measures were ranked at 2.1 (equivalent to USD100k), therefore it is deduced that they require the highest lifecycle cost among the proposed risk controls. The highest ranking according to some experts for the control measures exhibited values around 3 (which is equivalent to USD800k). Although these are approximate values, it provides an indication of scale of the required interventions, which

Table 10. Summary of risk control measures.

Item	Preventative measures	Mitigative measures
Inadequate situation awareness by the shipboard situation awareness system	Sensors fusion Testing More advanced sensors (information acquisition systems) Methods for training	Abnormalities' detections system Control transfer to RCC
Main and auxiliary power generation loss	Power/Take in/off Redundancy in power generation system components Self-reconfiguration of propulsion plant Preventative maintenance Intelligent monitoring of components	Emergency DG set Use of battery pack Automatic drop of anchor Visual and audible notification to ships
Thruster and steering system failures	Redundancy in components used for propulsion Self-reconfiguration for propulsion plant	Automatic drop of anchor, if propulsion completely unavailable Visual and audible notification to ships
Environment protection from oil pollution	(See other factors)	Self-closing valves in ventilation lines of fuel tanks Cofferdams (already implemented)
Cybersecurity related scenarios	Presence of antivirus on key controllers Double verification during software update	Remote rebooting Control transfer to RCC Contingency plans in place Safe shutdown procedures
Fire	Inerting using CO_2/N_2 during firefighting	Closing ventilation Visual and audible notification to ships Dropping anchor
Security related	Intrusion detection system Physical protection for the RCC	Control transfer to RCC

is a fraction of the initial ship capital cost. It must be noted that the provided figures are initial estimations, and it is expected that more accurate results will be obtained based on application of the novel technologies to the two demonstrators following the completion of the AUTOSHIP project.

Step 5: Identification of testing scenarios

For the hazards with the higher risk, in total 17 testing scenarios with associated pass/fail criteria were identified and listed in Table 11. The testing scenarios are related to: (a) performing the fail-safe functions; (b) relevant reconfiguration functions; and (c) the system effectiveness in specific scenarios.

For the navigation and situation awareness functions, it is proposed to implement testing in both the virtual and real environments prior to ship deployment, as well as to test the ability of the remote control centre to take over control in critical situations. In case where the ship loses communication with remote control centre, it is proposed to test the ship ability to sail to a safe location or to drop anchor depending on the situation. In the cases of power/propulsion loss, thrust loss, or electrical system failure, the following scenarios are proposed: (a) testing the system ability to reconfigure by using alternative power sources, connection types and thruster means; and (b) testing the safe anchoring procedure in the case of a complete power loss. For cases of fire emergencies, testing of the effective starting up and operation of the firefighting system, as well as the safe anchoring procedure are proposed.

Discussion on the process

As it was demonstrated in the preceding sections, the developed process is applicable during initial stages of the autonomous ships design, where only high-level information about the systems is available. Moreover, some initial design recommendations and testing requirements were developed based on the analysis results. In this way, the method can drive the decisions regarding ship design and verification early at the system design process, allowing more time for relevant testing arrangements to be planned and carried out.

An advantage of the proposed process is that it is interconnected with the FSA risk matrix and consequences classification. This is of great value for demonstrating the compliance of the initial design with the potential future maritime risk acceptance criteria to the relevant authorities. The proposed HAZID process employed a hybrid approach integrating both the ship functions and operational phases hierarchical structures as well as the use of specific guidance words. In this respect, the proposed method contributed to the systematisation and inclusiveness of the process, thus allowing for a more effective and thorough analysis.

Table 11. Critical hazards generic testing scenarios and generic fail/pass criteria.

Id	Hazard	Generic testing scenario (TS)	Fail/pass criteria
H32	The surrounding situation is not properly (wrongly or not) determined For example, objects are not detected and recognised (small objects, navigation marks, ships, ship lights, floating objects, depth, recreation crafts, people) weather is not properly conceived	A. Virtual testing of artificial intelligence (AI) algorithms used for image recognition in laboratory environment and during deployment under various environmental conditions with human presence B. Functional testing of system components under various environmental conditions C. Failure testing (testing of system performance when a component is faulty) in a simulated/real environment with human presence	1. Ship can effectively detect X% of objects in virtual and real environment (TS A). 2. Ship components function properly during various testing stages (unit testing, integration testing, sea trials) (TS B). 3. Ship can effectively detect X% of objects in virtual and real environment even if some components are faulty (TS C).
H13	Control loss over propulsion/thrust function	A. If ship can drop anchor in case of propeller loss	1. Ability of system to respond properly (TS A).
H33	Other ship not activating its AIS/lighting system/ communication system	A. Simulation testing of autonomous ship situation awareness system with ships operating that have switched off their AIS and communication system to ensure the functionality of sensor fusion system B. Sea trials for situation awareness behaviour when another ship has switched off its AIS and communication system C. Ship response to detected anomaly testing during HIL and sea trials	1. Ship can effectively detect X% of scenarios when the other ship has switched off their communication system and AIS in virtual ship (TS A). 2. Ship can detect the ship which has switched off its communication system and AIS system (TS B). 3. Ship can send an alarm to the RCC if it detects another ship with switched off communication system and AIS system (TS C). 4. Ship can transfer control to the RCC in the above case during sea trials and hardware in the loop testing (TS C).
H37	Ship on collision track with other ships in canal	A. Simulation testing of autonomous ship collision avoidance algorithm in virtual environment in a number of encountering conditions B. Sea trials for autonomous ship collision avoidance algorithm	1. Ship can effectively avoid X% of collision situations (TS A). 2. Ship does not collide in any situation (TS B).
H46	Ship losing communication with the remote control centre	A. Testing the ship ability to respond correctly to the communication loss during various testing (SIL, HIL, sea trials) B. Testing connectivity in the operating area	1. Ship responds as prescribed (it continues mission/navigates to a safe location/drops the anchor) (TS A). 2. Sufficient connectivity not impeding ship performance in the area (TS B).
H43	Loss of communication with other ships	A. If ship can reconfigure to another communication system. B. If ship can enter increase the distance from other ships in case of communication loss	1. Ability of system to reconfigure to another communication channel (TS A). 2. Ability of system to increase safety margin (TS B).
H36	Ship on collision track with bridge/lock	A. Simulation testing of autonomous ship passing through locks under various environmental and loading conditions B. Testing of situation awareness system when various lighting signals are provided from locks in virtual environment C. Sea trials for autonomous ship going though locks/bridges D. Testing of ship response when conflicting signals (e.g. the bridge is closed but the light is green) are provided in virtual environment	1. Ship can effectively avoid X% of contacts in simulated environment (TS A). 2. Ship can effectively and safely recognise X% of signals provided by locks in virtual environment (TS B). 3. Ship can safely pass from locks/ under bridge during sea trials (TS C). 4. Ship can recognise and report conflicting signals (TS D).

This is the main advantage of this approach compared to the classical HAZID methods, which are usually based on what-if process and simple brainstorming sessions. The proposed steps can lead to the semi-automation of the HAZID process and development of corresponding software tools to facilitate the more effective HAZID process implementation.

During the process, it was found that the experts rankings can exhibit high variance, which resulted in the high rankings uncertainty. This uncertainty was accounted for evaluating the hazardous scenarios and defining their criticality. This is in line with the risk definition as impact of uncertainty on the outcome.[33] Future studies can be benefitted by including statistical data from the semi-automated ships, where full or partial automation is already in place. This would allow for the gradual accumulation of relevant data, thus rendering the analysis results more effective. Employing weights in the individual experts rankings (typically affected by each expert' experience) will not result in the rankings uncertainty reduction, as the uncertainty is already high.

The HAZID workshops were implemented online due to the COVID-19 restrictions. This contributed to a number of challenges, as it was not straight-forward to establish a functional discussion protocol. Slight communication gap issues were identified and resolved, which influenced the process of updating the hazard list, the ranking process and review process or risk control measures, leading to some experts not ranking some hazards or not sufficiently realising their context. It can be argued that the use of a structured approach, the incorporation of previous analyses, the distribution of the material in advance, the thorough discussions on the design and use of the rankings uncertainty were implemented as counter-measures to mitigate the communication gap to the extent possible. Yet, some residual risk not appropriately addressed is expected due to the online procedures. As a recommendation for future online application of the process, the facilitator could consider the distribution of relevant material in advance, breaking down the online session in a number of smaller sessions, starting the workshops discussion with the scenarios that he/she would consider as critical, dedicating more time to the discussion with respect to the rankings and allocating sufficient time to clarify ambiguities with respect to the initial design. It is also recommended to include a diverse group of experts from safety, security, cybersecurity disciplines as well representing variety of stakeholders (researchers, authorities, operator, original equipment manufacturers).

Compared to other safety methods, it is commented that the process is not competitive to the STPA, FMEA, FTA, CASA,[63] etc. but rather complimentary. Based on the followed approach, the critical hazards were identified. At a subsequent phase, STPA can be used for identifying the relevant Unsafe Control Actions and their causal factors for these critical hazards. The causes identified in this study can be re-employed. Similarly, the proposed approach can highlight the critical hazards, thus facilitating their further analysis by employing other safety methods, such as FTA, ETA and FMEA. A dedicated cybersecurity risk assessment is also required to reduce uncertainty, as for instance reported in Bolbot et al.[9] or in BV[64] guidance.

Due to the limitations of the employed HAZID process, the identified testing scenarios are generic, hence they need to be developed further into specific test cases by adding more specific details on test acceptance criteria. To resolve this issue, other methods need to be employed. In addition, more information is required to specify coverage criteria during the testing procedures. However, the advantage of employing the proposed HAZID process is that these high-level testing scenarios can be identified early during the autonomous ship system design process guiding the process of designing test scenarios and increasing the project slack.

Further development of the proposed risk analysis process could focus on its automation and standardisation which would allow comparison of various results of risk assessment. The presented process was developed for the case study of the IWW autonomous ships; however, it can be applied to other ships including Short Sea Shipping and ocean-going ships. In these cases, the customisation of the relevant acceptance criteria is required.

Conclusions

In this study, a novel hybrid, semi-structured hazard identification (HAZID) process for the risk assessment of autonomous ships that employs operational and functional hierarchical categorisations was presented. The process integrates safety, security and cybersecurity analyses at the initial design stage of the autonomous ships. The process was applied for the risk assessment, identification of risk control measures and the development of testing scenarios for the preliminary design of the theoretical unmanned IWW ship.

The main findings of this study are as follows:

- The proposed semi-structured HAZID process supported the identification of relevant hazardous scenarios in a systematic way by incorporation of ship functions and operational phases, whilst comprehensively considering hazards due to safety, security and cybersecurity issues.
- The process supported the ranking of the scenarios and identification of the critical ones by considering uncertainty.
- Effective communication is crucial for the robust HAZID process application.
- More than 80 hazardous scenarios were identified for the investigated IWW autonomous ship.

- A number of hazards and consequences are similar for the manned and unmanned IWW ships, yet their causes are different.
- The reputational and third-parties safety risks were highlighted as the one with the highest severity for the IWW ship.
- The contribution of safety related causes was found of greater impact compared to the cybersecurity and security related causes, but the contribution of cyber security or security causes should not be ignored.
- Based on the risk assessment results, it can be anticipated that increased automation level on small ships can proceed faster due to reduced risks compared to large ships. However, special attention should be paid to navigational risk.
- Uncertainty during the rankings was found considerable, especially for the frequency ranking, due to the lack of pertinent statistical data and expertise in the novel technologies that have been under development.
- The scenarios related to navigation functions were found to be critical for the investigated IWW ship due to high RI index value and the associated uncertainty.
- Due to the high uncertainty in the frequency ranking, the mitigation measures can be considered as more effective to confine the risk. This does not mean that the preventative measure should be ignored.
- It was estimated that the cost of the required measures is a fraction of the cost of the ship; however more detailed analyses are required to verify these cost results.

The proposed HAZID and risk assessment process can also be applied to other autonomous and remotely controlled ships for the identification of risks and development of relevant safeguards. A future research could focus on the automation and standardisation of the risk assessment approach across the maritime community.

Acknowledgements

The authors kindly acknowledge the comments and input with respect to requirements provided from Kongsberg Maritime. The authors affiliated with the MSRC greatly acknowledge the funding from DNV AS and RCCL for the MSRC establishment and operation. The opinions expressed herein are those of the authors and should not be construed to reflect the views of DNV AS, RCCL, Bureau Veritas, DVW, SINTEF, Kongsberg Maritime, Zulu associates or the acknowledged individuals and their associated organisations.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was carried out in the framework of the AUTOSHIP project, which is funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 815012.

ORCID iDs

Victor Bolbot https://orcid.org/0000-0002-1883-3604
Gerasimos Theotokatos https://orcid.org/0000-0003-3547-8867
Lars Andreas Wennersberg https://orcid.org/0000-0002-4090-3699

References

1. Autoship. Autonomous shipping initiative for European waters, https://www.autoship-project.eu/ (2019, accessed 01 February 2021).
2. Bolbot V, Theotokatos G, Bujorianu LM, et al. Vulnerabilities and safety assurance methods in cyber-physical systems: a comprehensive review. *Reliab Eng Syst Saf* 2019; 182: 179–193.
3. Eloranta S and Whitehead A. Safety aspects of autonomous ships. In: *6th international maritime conference on design for safety* (ed Gl DNV), Hamburg, Germany, 28-30 November 2016, pp.168–175. DNV GL
4. Kavallieratos G, Katsikas S, Gkioulos V, et al. Cyber-attacks against the autonomous ship. In: Katsikas SK, Cuppens F and Cuppens N (eds) *Computer security*. Cham: Springer International Publishing, 2019, pp.20–36.
5. Kavallieratos G, Katsikas S and Gkioulos V. Cybersecurity and safety Co-engineering of cyberphysical systems—A comprehensive survey. *Future Internet* 2020; 12: 65.
6. Wingrove M. Shipborne systems most vulnerable to cyber-attack. *Mar Electron Commun* 2017; 11: 27.
7. NTSB. *Collision between US Navy destroyer fitzgerald and philippine-flag container ship ACX crystal Sagami Nada Bay off Izu Peninsula, Honshu Island*, Washington, DC: National Transportation Safety Board, 2020.
8. IMarEST. Close encounter with a nuclear submarine, http://think.imarest.org/q/17IwQ8DXgjxs4RVPaXO-GiC/wv, (2020, accessed 30 September 2020).
9. Bolbot V, Theotokatos G, Boulougouris E, et al. A novel cyber-risk assessment method for ship systems. *Saf Sci* 2020; 131: 104908.
10. Accident Investigation Board Norway. *Interim report on the investigation into the loss of propulsion and near grounding of Viking sky*. Lillestrøm, Norway: AIBN, 2019.
11. Veritas Bureau. Guidelines for autonomous shipping. In: Veritas B (ed.). *NI 641 DT R01E*. Paris: Bureau Veritas, 2019.
12. DNV GL. Autonomous and remotely operated ships. GL D, (ed.). *DNVGL-CG-0264*, 2018, https://rules.dnv.com/docs/pdf/DNV/cg/2018-09/dnvgl-cg-0264.pdf (accessed 01 February 2020)
13. IMO. Interim guidelines for MASS trials. *MSC1/Circ1604*, 2019, https://www.register-iri.com/wp-content/uploads/MSC.1-Circ.1604.pdf (accessed 01 february 2020)
14. EMSA. *Study of the risks and regulatory issues of specific cases of MASS*, 2020, http://emsa.europa.eu/about/items.html?cid=2&id=3892 (accessed 01 June 2020)

15. Kretschmann L, Rødseth Ø, Tjora Å, et al. *Qualitative assessment*, 2015, http://www.unmanned-ship.org/munin/wp-content/uploads/2015/10/MUNIN-D9-2-Qualitative-assessment-CML-final.pdf (accessed 01 June 2019)
16. Rødseth ØJ and Burmeister H-C. Risk assessment for an unmanned merchant ship. *TransNav Int J Mar Navig Saf Sea Transp* 2015; 9: 357–364.
17. Thieme CA, Guo C, Utne IB, et al. Preliminary hazard analysis of a small harbor passenger ferry – results, challenges and further work. *J Phys Conf Ser* 2019; 1357: 012024.
18. Wróbel K, Montewka J and Kujala P. Towards the assessment of potential impact of unmanned vessels on maritime transportation safety. *Reliab Eng Syst Saf* 2017; 165: 155–169.
19. Valdez Banda OA and Goerlandt F. A STAMP-based approach for designing maritime safety management systems. *Saf Sci* 2018; 109: 109–129.
20. Utne IB, Rokseth B, Sørensen AJ, et al. Towards supervisory risk control of autonomous ships. *Reliab Eng Syst Saf* 2020; 196: 106757.
21. Wróbel K, Montewka J and Kujala P. Towards the development of a system-theoretic model for safety assessment of autonomous merchant vessels. *Reliab Eng Syst Saf* 2018; 178: 209–224.
22. Rokseth B, Haugen OI and Utne IB. Safety verification for autonomous ships. In: *MATEC web of conferences*, 2019, p.02002. EDP Sciences. https://doi.org/10.1051/matecconf/201927302002.
23. Wrobel K, Krata P and Montewka J. Preliminary results of a system-theoretic assessment of maritime autonomous surface ships' safety. *TransNav Int J Mar Navig Saf Sea Transp* 2019; 13: 717–723.
24. Ventikos NP, Chmurski A and Louzis K. A systems-based application for autonomous vessels safety: hazard identification as a function of increasing autonomy levels. *Saf Sci* 2020; 131: 104919.
25. Glomsrud JA and Xie J. A structured STPA safety and security co-analysis framework for autonomous ships. In: *European safety and reliability conference* (eds. M Beer and E Zio), Hannover, Germany, 22-26 September 2019. Research publishing.
26. Zhou X-Y, Liu Z-J, Wang F-W, et al. Towards applicability evaluation of hazard analysis methods for autonomous ships. *Ocean Eng* 2020; 214: 107773.
27. Chaal M, Valdez Banda OA, Glomsrud JA, et al. A framework to model the STPA hierarchical control structure of an autonomous ship. *Saf Sci* 2020; 132: 104939.
28. Kavallieratos G, Katsikas S and Gkioulos V. SafeSec Tropos: joint security and safety requirements elicitation. *Comput Stand Interfaces* 2020; 70: 103429.
29. SAE. ARP4761. *Guidance and methods for conducting the safety assessment process on civil ariborn systems and equipment*. Warendale, PA, USA: SAE, 1996.
30. PD ISO/PAS 21448: Road vehicles – safety of the intended functionality. London, UK: British standards institute, 2019.
31. ISO 10218. Robots and robotic devices—safety requirements for industrial robots, London, UK: British standards institute, 2011, pp.1-2.
32. ISO 31010:2009. Risk management—risk assessment techniques. Geneva, Switzerland: International Organization for Standardization, 92.
33. ISO 31000:2018. Risk management – guidelines. London: British Standards Institution.
34. IMO. *MSC. 1/circ 1455 guidelines for the approval of alternatives and equivalents as provided for in various IMO instruments*. London: IMO, 2013.
35. UK M. Maritime autonomous surface ships UK code of practice, 2018, https://www.maritimeuk.org/media-centre/publications/maritime-autonomous-surface-ships-uk-code-practice/ (accessed 01 April 2020).
36. LR. LR code for unmanned marine systems. LR, (ed.), 2017, https://www.lr.org/en/unmanned-code/ (accessed 01 August 2020).
37. Kriaa S, Pietre-Cambacedes L, Bouissou M, et al. A survey of approaches combining safety and security for industrial control systems. *Reliab Eng Syst Saf* 2015; 139: 156–178.
38. International Maritime Organisation. *Revised guidelines for formal safety assessment (FSA) for use in the IMO rule-making process*. London: IMO, 2018.
39. IMO. *Revised guidelines for formal safety assessment (FSA) for use in the IMO rule-making process*. London: IMO: IMO, 2018.
40. Leveson N and Thomas J. *STPA handbook*, https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf (2018, accessed 01 September 2019)
41. Ahluwaja A. *Managing new technology risks – DNV GL technology qualification process*, https://www.aiche.org/sites/default/files/community/291721/aiche-community-site-page/315581/technologyqualificationpresaicheamar-ahluwalia8-3-18.pdf (2018, accessed 01 August 2020)
42. EMSA. *Study on electrical energy storage for ships*, http://www.emsa.europa.eu/publications/reports/item/3895-study-on-electrical-energy-storage-for-ships.html (2020, accessed 01 June 2020).
43. Bruce BF. *The SAGE encyclopedia of educational research, measurement, and evaluation*. Thousand Oaks, California: SAGE Publications, 2018.
44. Kaplan S and Garrick BJ. On the quantitative definition of risk. *Risk Anal* 1981; 1: 11–27.
45. Bureau Veritas. Risk based qualification of new technology – methodological guidelines. *NI 525 DT R01 E*, https://marine-offshore.bureauveritas.com/ni525-risk-based-qualification-new-technology-methodological-guidelines (2020, accessed 01 September 2020).
46. Flage R and Aven T. Expressing and communicating uncertainty in relation to quantitative risk analysis. *Reliability: Theory & Applications* 2009; 4: 9–18.
47. Goerlandt F and Reniers G. On the assessment of uncertainty in risk diagrams. *Saf Sci* 2016; 84: 67–77.
48. Aerts A, Reniers M and Mousavi MR. Model-Based testing of cyber-physical systems. In: Rawat DB, Jeschke S and Brecher C (eds) *Cyber-physical systems*. Boston: Academic Press, 2017, pp.287–304.
49. Asadollah SA, Inam R and Hansson H. A survey on testing for cyber physical system. In: *Proceedings of the 27th IFIP WG 6.1 International Conference ICTSS*, Sharjah and Dubai, United Arab Emirates, 23-25 November 2015, pp.194–207. Springer.
50. Rokseth B, Utne IB and Vinnem JE. Deriving verification objectives and scenarios for maritime systems using the systems-theoretic process analysis. *Reliab Eng Syst Saf* 2018; 169: 18–31.
51. Rødseth and Burmeister H-C. *New ship designs for autonomous vessels*, 2015. MUNIN project.
52. Bolbot V, Theotokatos G, Boulougouris E, et al. Safety related cyber-attacks identification and assessment for autonomous inland ships. In: *Internation seminar on*

safety and security of autonomous vessels (ISSAV), Helsinki, Finland, 17-20 September 2019. Helsinki, Finland: Aalto University.
53. Höyhtyä M, Huusko J, Kiviranta M, et al. Connectivity for autonomous ships: Architecture, use cases, and research challenges. In: *2017 international conference on information and communication technology convergence (ICTC)*, Jeju, South Korea, 18–20 October 2017, pp.345–350. New York, NY: IEEE.
54. van Cappelle LE, Chen L and Negenborn RR. Survey on short-term technology developments and readiness levels for autonomous shipping. *Computational logistics*. In: *Proceedings of the 9th International Conference, ICCL* (eds Cerulli R, Raiconi A, Voß S, et al, Vietri sul Mare, Italy, 1–3 October 2018, pp.106–123. Springer International Publishing.
55. Geertsma RD, Negenborn RR, Visser K, et al. Design and control of hybrid power and propulsion systems for smart ships: a review of developments. *Appl Energy* 2017; 194: 30–54.
56. Chaal M, Banda OV, Glomsrud Jon Arne, et al. A framework to model the STPA hierarchical control structure of an autonomous ship. *Safety Science* 2020; 132: 104939.
57. Wennersberg LA and Nordahl H. *D2.1 – complete supply chain mapping & identifications of interactions between SSS and IWW demonstrators*, 2019.
58. Blue Lines Logistics. Blue Lines logistics news, https://zulu-associates.com/ (2021, accessed 01 February 2021).
59. IMO. Regulatory scoping exercise, https://www.imo.org/en/MediaCentre/HotTopics/Pages/Autonomous-shipping.aspx (2020, accessed 01 November 2020).
60. Central Commission for the Navigation of the Rhine. Definitions on various forms of automated navigation, 2018.
61. Rødseth ØJ, Faivre J, Hjørungnes SR, et al. *AUTOSHIP deliverable D3.1 autonomous ship design standards, revision 2.0*, 2020.
62. Puisa R, Lin L, Bolbot V, et al. Unravelling causal factors of maritime incidents and accidents. *Saf Sci* 2018; 110: 124–141.
63. Bolbot V, Theotokatos G, Boulougouris E, et al. A novel method for safety analysis of cyber-physical systems—application to a ship exhaust gas scrubber system. *Safety* 2020; 6: 26.
64. BV. Rules on cyber security for the classification of marine units. BV, (ed.). *NR 659 DT R01*. Paris, France, 2020.

Appendix A

Abbreviation and notation

Abbreviation	Definition
BV	Bureau veritas
FI	Frequency index
FSA	Formal safety assessment
HAZID	Hazard identification
IMO	International maritime organisation
ISO	International standard organisation
MASS	Maritime autonomous surface ships
RI	Risk index
SI	Severity index
STPA	System-theoretic process analysis
UL	Uncertainty level

Appendix B

Detailed list of hazards

The detailed list of hazards identified through the proposed process is provided in Table A1.

Table A1. Hazard list.

Operating phase	IWW use case hazards/inadvertent events/(*guidewords*)
Mission planning	Selecting route with heavy traffic/(*wrong output*) Selecting route with bad weather conditions/ too shallow water/low visibility/ice in water/(*wrong output*) Selecting route with too low bridge/too narrow canals/ (*wrong output/too little*) Hackers stealing the intended path/route/(*steal*) Excessive load of ship metallic structure due to improper loading/ (*wrong output/too much*) Fuel/lubricants are not enough to accomplish the mission/(*wrong output/too little*) Wrong/No equipment health predictions resulting in improper management of systems maintenance/ (*wrong/no output*)
Activities in ports	Terrorists taking hostage of the technical personnel and sailing/guarding themselves on the ship/ (*hostage*) Ship getting unmoored/(*untimely output*) Damage to situation awareness equipment due to the interacting object, port cranes, etc.)/(*wrong input*) Loss of communication/(*no output*) Ship drifting away/(*untimely output*) Malware transferred to ship (and RCC) through port facilities/(*transfer*) Violating stability criteria during loading and offloading/(*wrong output*)

(continued)

Continued

Operating phase	IWW use case hazards/inadvertent events/(*guidewords*)
	Malware transferred from ship (and RCC) to port facilities/(*transfer*)
	Loss of auxiliary power/(*no output*)
	Hooligans attacking the ship/(*attack*)
	Disturbances in the electric distribution systems/(*wrong output*)
	Loss of auxiliary power due to shore connection issues/(*no output*)
Arrival in ports/ departure from ports	Loss of main power/(*no output*)
	Control loss over propulsion/thrust function/(*conflicting output*)
	Total loss of auxiliary power/(*no output*)
	Disturbances in the electric distribution systems/(*wrong output*)
	Control loss over steering equipment/(*conflicting output*)
	Ship position unstable during approaching the harbour/(*conflicting output*)
	Ship transmitting information to wrong actors/(*get control over*)
	Lower propulsion power/or slower propulsion power is provided than required resulting in loss of manoeuvrability/(*wrong untimely output*)
	Leaving with irrelevant personnel, for example, port, pilots, technicians onboard/(*wrong output*)
	Ship stuck in port/(*no output*)
	No route selection/Not starting from the quay/port/lock/(*no output*)
Transit	Other ship not activating its AIS/lighting system/communication system/(*no output*)
	Ship on collision track with other ships in canal/(*wrong output/too much*)
	The surrounding situation is not properly (wrongly or not) determined, for example, Objects are not detected and recognised (small objects, navigation marks, ships, ship lights, floating objects, depth, recreation crafts, people) weather is not properly conceived/(*wrong/no output*)
	Ship on collision track with bridge/lock (*wrong output/ too much*)
	Ship losing communication with the RCC/(*no output*)
	Control loss over propulsion/thrust function (Thruster/propeller not starting or stopping during operation)/(*no/untimely/conflicting output*)
	Control loss of main power/(*no/untimely/conflicting output*)
	Ship can be on collision track with floating/submerged objects, for example, submerged car, submarine, bike/plastic/(*wrong output/too much*)
	Loss of communication with other ships/(*no output*)
	Loss of auxiliary power/(*no output*)
	Disturbances in the electric distribution systems to essential consumers (thrusters, main engine auxiliary systems, etc.)/(*wrong/no output*)
	Control loss over steering equipment/(*no output*)
	Violating stability criteria during travel/(*wrong output*)
	Ship transmitting erroneous information to the other ships endangering the safety of navigation/(*get control over*)
	Ship unable to go through lock/under bridge/(*wrong output*)
	Ship blocking access to lock/bridge passage/(*wrong output*)
	Loss of communication with lock/bridge/(*untimely function output*)
	Ship navigating to other destination port, than desired due to piracy act/smuggler/(*hijacking/stealing*)
	Bridge is lowered while barge is passing/(*conflicting output*)
	Stealing images from the video cameras (any type of hackers)/ (*steal*)
Emergency	Inadequate firefighting during fire/(*no/wrong output*)
	Remote control centre not taking over the operations in critical situation/(*no output*)
	Loss of control over emergency power (when needed)/(*no output*)
	Inadequate power generation during fire/flooding/(*wrong/no output*)
	Ship unable to depart from the quay during accident in port/lock/(*no output*)
	Excessive load of ship metallic structure due to contact/(*wrong input*)
	Excessive load of ship metallic structure due to fire/(*wrong input*)
	Excessive load of ship metallic structure due to collision/(*wrong input*)
	Excessive load of ship metallic structure due to grounding/(*wrong input*)
	Excessive load of ship metallic structure due to explosion/(*wrong input*)
	Water ingression is not identified/(*no output*)
	No Safe and Rescue operation/(*no output*)
	Loss of damage stability control during flooding/(*no output*)
Maintenance and repairs	Excessive corrosion of metallic structure/(*wrong input*)
	Not implementing maintenance of critical systems (propulsion, power, navigation), which are due to fail in the next voyage/(*no input*)
	Not implementing maintenance of electronic systems responsible for control of critical functions (propulsion, power, navigation)/(*no input*)
Any mode	Cyber-attack on ship systems (referencing to all the scenario related to cyberattacks on ship systems mentioned in the excel, below and above)/(*not applicable*)

(continued)

Continued

Operating phase	IWW use case hazards/inadvertent events/(*guidewords*)
	No safety monitoring/(*no output*) Fuel/lubricants leakage (inflammable medium leakage)/(*wrong output*) Trafficking drugs/weapons on the ship/(*smuggling*) Overheating of nearby equipment/surfaces/(*wrong output (too much)*) Flooding caused by ship systems/(*conflicting control action*) Hackers leaking critical ship voyage information/(*steal*) Arson (deliberate attempt from malicious insiders to put on fire)/(*damage*) Cyber terrorists attempting ignition by causing fuel leakage and overheating/(*damage*) Uncontrolled short circuits/(*conflicting output*) False firefighting equipment operation when is not required/(*untimely output*) Arcs in switchboards/(*wrong output*) Stealing the condition monitoring data/digital model through communication links (competitors)/(*steal*) Excessive oil emission to the canals/(*wrong output (too much)*) Stealing the condition monitoring data/digital twin model by boarding the ship (competitors)/(*steal*) Hackers destroying the ship certificates and critical information/(*destroy*) Hackers leaking the ship certificates and critical information/(*steal*)