Ocean Engineering 251 (2022) 111100

Contents lists available at ScienceDirect

Ocean Engineering

journal homepage: www.elsevier.com/locate/oceaneng

Towards an online risk model for autonomous marine systems (AMS)

Ruochen Yang[*], Ingrid Bouwer Utne

Department of Marine Technology, Norwegian University of Science and Technology (NTNU), Otto Nielsens veg 10, Trondheim, 7491, Norway

ARTICLE INFO

Keywords:
Maritime system safety
Online risk model
Autonomous marine systems (AMS)
STPA
PHA
HAZOP

ABSTRACT

Considering that few or no human operators are directly involved in the operation of Autonomous Marine Systems (AMS), an online risk model is necessary to enhance the intelligence of the AMS, its situation awareness, and decision-making. The current study identifies the criteria for an online risk model for AMS, which can be used to assess its validity and effectiveness.

Taking an under-ice Autonomous Underwater Vehicle (AUV) operation as an example, the current work investigates how different risk analysis methods, namely the Preliminary Hazard Analysis (PHA), the Systems Theoretic Process Analysis (STPA), and Procedural Hazard and Operability Analysis (HAZOP), contribute to fulfilling the different criteria for online risk modeling of AMS. The analysis results show that STPA can be considered a good basis for developing an online risk model due to its relatively good coverage of the identified evaluation criteria, especially its ability to handle the interaction between system and software failure. In addition, considering some shortcomings of using STPA and the changing role of human operators in the AMS operation, PHA and Procedural HAZOP can be used as complementary tools. It is expected that the analysis results and conclusions can be adapted to other AMS as well.

1. Introduction

The development of Autonomous Marine Systems (AMS), including Marine Autonomous Surface Ships (MASS), Unmanned Underwater Vehicles (UUV), and autonomous offshore oil and gas systems is emerging due to the potential for improved safety and efficiency. Compared to conventional marine systems, AMS are expected to operate with few or even no crew onboard in the future, and it is therefore essential to ensure that AMS have the expected level of reliability, availability, maintainability and safety to be acceptable for widespread use at sea. At the very least, AMS should be as safe as conventional marine systems (Laurinen, 2016). Hence, risk assessment is a necessary tool for the safe operation of AMS and to provide information for decision-makers, including both operators and the AMS itself.

Several previous studies have been conducted focusing on risk aspects of AMS. Chaal et al. (2020) proposed a framework for the Systems Theoretic Process Analysis (STPA) and its hierarchical control structure of an autonomous ship by making use of the knowledge gained in traditional ship operation, assuming that automated controllers will replace human controllers. Wróbel et al. (2017, 2018a, b) established a possible safety control structure for autonomous ships and conducted safety analysis to provide design recommendations for autonomous ships in terms of regulations, organization, and technology. Thieme et al. (2018) assessed the applicability of several existing ship risk models to MASS. The results demonstrate that, with extra consideration of the aspects of software and control algorithms and human-machine interaction, some existing risk models might be used as a basis for developing relevant risk models for MASS. Thieme and Utne (2017) proposed a process for developing safety indicators for the operation of AMS, reflecting the safety aspects of AMS operation to assist in operational planning, daily operational decision-making, and identification of improvements.

Several risk-related studies have been conducted specifically for UUV. Utne and Schjølberg (2014) proposed a taxonomy for hazardous events. Further, the results demonstrated that the main risk to humans in Autonomous Underwater Vehicle (AUV) operations in Arctic areas is during the launch and recovery of the vehicle. In a study by Brito and Griffiths (2011), a Markov chain model was applied to assess the reliability of AUVs, capturing the different states of the AUV operation. Step sequences from prelaunch to operation to recovery were included in this study, and a total of 11 discrete states were identified. A case study using the fault history of the Autosub3 AUV was conducted to provide the information for different operational phases. In another study by Brito and Griffiths (2016), the Bayesian approach was used to predict the risk of AUV loss during their missions. This research provided a rigorous

[*] Corresponding author.
E-mail address: ruochen.yang@ntnu.no (R. Yang).

https://doi.org/10.1016/j.oceaneng.2022.111100
Received 4 July 2021; Received in revised form 12 February 2022; Accepted 12 March 2022
Available online 25 March 2022

Abbreviation

ADCP	Acoustic Doppler Current Profiler
AMS	Autonomous Marine Systems
BBN	Bayesian Belief Network
BT	Bow Tie
CMBRA	Condition Monitoring-Based Risk Assessment
CTD	Conductivity, Temperature, Depth
DBN	Dynamic Bayesian Network
DRA	Dynamic Risk Assessment
DVL	Doppler Velocity Logs
ETA	Event Tree Analysis
FPSO	Floating Production Storage Offloading
FTA	Fault Tree Analysis
GNSS	Global Navigation Satellite System
HAZOP	Procedural Hazard and Operability Analysis
HCL	Hybrid Causal Logic
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
LBL	Long Baseline
LoA	Levels of Autonomy
MASS	Marine Autonomous Surface Ships
OH&S	Occupational Health and Safety
PHA	Preliminary Hazard Analysis
RIFs	Risk Influencing Factors
ROV	Remotely Operated Vehicles
STAMP	System-Theoretic Accident Model and Processes
STN	Single Transponder Navigation
STPA	Systems Theoretic Process Analysis
SVA	Security Vulnerability Analysis
UCAs	Unsafe Control Actions
UUV	Unmanned Underwater Vehicles
VHF	Very High Frequency
VTS	Vessel Traffic Service

procedure for AUV risk management in hazardous environments.

Loh et al. (2019) conducted a risk assessment for AUV under-ice missions to explore the risk of AUV missions in a harsh environment. Historical fault log data, as well as expert knowledge, were used in this study to develop a risk model. More studies on the risk analysis of AUV operations can be found in the review article by Chen et al. (2021). Hegde et al. (2018a, 2019) developed dynamic safety envelopes for autonomous Remotely Operated Vehicles (ROV). In these studies, the Octree method was used to set up the cuboidal shape of the proposed safety envelope, while the size of the dynamic safety envelope was determined by modeling a fuzzy inference system.

With few or no operators, an AMS needs improved perception, situation awareness, and planning/re-planning capabilities compared to the conventional marine system. For safe operation of the AMS, risk should be an essential factor that needs to be monitored and taken into account for control action. Therefore, an online risk model that is able to assess the possible risk dynamically and support the decision-making of the AMS is necessary (Utne et al., 2020). Few works, however, have been conducted to identify the specific needs for an online risk model of AMS and to analyze the applicability of the existing methods.

The current study identifies criteria for online risk models for AMS, using the systems engineering process. The identified evaluation criteria reflect the aspects that should be considered and included when developing an online risk model for AMS. In the paper, an AUV is used to investigate how the different existing risk analysis methods, i.e., Preliminary Hazard Analysis (PHA), STPA, and Procedural Hazard and Operability Analysis (HAZOP), contribute to fulfilling the criteria for online risk modeling of AMS. Further, the results from the analyses are evaluated with respect to developing an online risk model. Since the criteria are more or less generic, it is expected that the analysis results and conclusion could be adapted to other AMS as well.

The paper is structured as follows: Section 2 demonstrates the need for online risk modeling of AMS. The criteria for the assessment of the online risk models for AMS are identified in Section 3. In Section 4, some existing methods that might be used as a basis for the online risk models are briefly introduced. Risk analyses of an AUV under-ice operation using PHA, STPA, and Procedural HAZOP are performed as a case study in Sections 5 and 6. Section 7 summarizes the results from three methods for improved engineering design, operational procedures and further research, and investigates how the different analyses contribute to fulfilling the criteria for an online risk model of AMS. Section 8 concludes the current study and analyzes how the results from the analyses can be used to develop an online risk model.

2. On the need for online risk modeling of AMS

A risk model is a qualitative or quantitative representation of a system, measuring its risk level. In order to accurately measure the risk level, risk models are developed to capture the interaction between subsystems or events based on risk analysis. A typical risk analysis tries to answer three main questions (Rausand, 2013): (1) What can go wrong? (2) What is the likelihood of that happening? and (3) What are the consequences? Different types of risk analysis have been developed in the past decades and with different advantages and disadvantages, and they have been applied in a wide range of fields in both research and industry.

Traditional risk methods and models, such as Fault Tree Analysis (FTA) and Event Tree Analysis (ETA), usually provide a static risk picture of a system or an operation based on historical data or expert knowledge, but cannot capture the change of the system's risk level, which may deteriorate with time due to natural and management causes. In order to deal with the possible time-varying risk level, the concept of Dynamic Risk Assessment (DRA) was proposed, which aims to "update estimated risk of a deteriorating process according to the performance of the control system, safety barriers, inspection and maintenance activities, the human factor, and procedures" (Khan et al., 2016).

Several studies have been conducted to address the dynamic risk model in the past decade to take advantage of updated information, especially using Bayes' theorem (Baksh et al., 2018; Barua et al., 2016; Khakzad et al., 2012, 2013; Liu et al., 2021; Paltrinieri et al., 2014; Wang et al., 2017; Yang et al., 2020a). Khakzad et al. (2012) proposed an updated Bow Tie (BT) method to achieve dynamic risk assessment by updating safety barriers of BT using Bayes' theorem. The prior failure rate of each safety barrier is assumed to follow a gamma distribution. The number of failures over time is taken into account to form likelihood functions, which is then used to update the failure rate estimation using Bayes' theorem. A novel Bayesian Belief Network (BBN) framework was developed by Baksh et al. (2018) to model marine transportation accidents in Arctic waters. The model is capable of updating the results whenever new evidence is available during the operation using Bayes' theorem.

Barua et al. (2016) developed a dynamic operational risk assessment method for the chemical process industries, which takes into account the sequential dependency and the effect of time. The changes of variables over time are represented as the temporal dependencies between two discrete time slices using conditional probability in a Dynamic Bayesian Network (DBN). Several studies used a similar approach in other fields, such as fire accidents (Wang et al., 2017) and AUV operations (Yang et al., 2020a).

Most of the existing DRA methods, however, rely on incident/accident statistics or temporal dependence based on historical or experience data to update the risk estimation, which means that they must wait until accidents or near misses occur before updating the estimation of the risk indexes (Zio, 2018). Therefore, these methods may fail to reflect the rapid changes of the operating environment and system status and provide timely support for decision-making during the operation of AMS. The development of wireless technology, cheaper and more advanced sensor technology, and improved computational capability are promoting the development of a more dynamic and online risk assessment (Vinnem et al., 2015; Zio, 2018).

The concept of online risk management was first proposed by Vinnem et al. (2015), in which the online risk models are built on data from different sources, including historical data, sensors and measurements, and experience data. With the help of appropriate data interpretation methods, online risk models provide pre-warnings of possible operational deviations. Though the framework was first proposed for Floating Production Storage Offloading (FPSO), the concept has in recent years been used in other fields, such as for autonomous ships (Utne et al., 2020). A similar idea to the online risk model was also proposed by Zio (2018), which is called Condition Monitoring-Based Risk Assessment (CMBRA). While most existing DRA methods rely on statistical data for risk estimation to update risk, the proposed CMBRA enables the risk estimation to be updated by using condition-monitoring data.

Utne et al. (2020) outlined a framework for online risk modeling for an autonomous ship. The hazard identification is conducted using the STPA, and the results are used to develop a BBN risk model, in which sensor data can be used to measure monitorable variables as part of the autonomous ship's supervisory risk control. Zeng and Zio's (2018) work presents a dynamic risk assessment method, combining statistical and condition-monitoring data, that allows for the estimation of risk based on data collection during operation. A BBN model with simulations is developed to utilize two types of data: statistical data provides the historical information about the system, while condition-monitoring data provides the degradation status of the specific target system and describes system-specific features. Several other studies also attempt to make use of condition-monitoring data in risk assessment (Kim et al., 2015; Lazakis et al., 2016; Zadakbar et al., 2015).

3. Evaluation criteria for the online risk model of AMS

To identify relevant criteria for the online risk model of AMS, a system engineering process is used, based on (Blanchard, 2004). The functional requirements for AMS with respect to risk and online risk models are described. The requirements identified are then used to derive the evaluation criteria, which reflect aspects that should be represented in an online risk model for AMS. The purpose is to identify potential gaps and focus areas that need to be especially addressed when developing online risk models for AMS. Furthermore, the purpose is to assess the efficiency of existing risk methods as a basis for such models.

3.1. Functional requirements

Table 1
Requirements for AMS with respect to risk, adapted from Thieme et al. (2018).

Requirements	Description
R1.1	Reliable and timely identification of hazards and hazardous events
R1.2	Reliable and verified hardware during operation (sensors, machinery, and control system)
R1.3	Reliable and verified software and algorithms and software updates during operation
R1.4	Robust interaction between software and hardware
R1.5	Reliable and adequate communication/interaction between AMS (includes crew if any) and the external supporting system (if any)
R1.6	Reliable and adequate communication between AMS and other marine stakeholders
R1.7	Reliable and adequate provisions for adaptive autonomy/mode
R1.8	Accessible and affordable human–machine interfaces

Table 1 summarizes the functional requirements of AMS with respect to risk. This table is adapted from the work of Thieme et al. (2018), expanding the scope from MASS only to online risk modeling of other AMS, such as UUVs, and autonomous offshore platforms. During the operation, the AMS should identify in a timely way the potential hazards and hazardous events (R1.1), supporting the decision-making and risk control by either operators or the system itself. Hardware, such as machinery, sensors, and the control system, need to perform their desired function during the operation (R1.2). Compared to conventional marine systems, the software and algorithms involved in AMS will increase. Issues due to the introduction or increase of the software in AMS should be solved. The software and algorithms should execute their functions in a reliable and safe manner and be verified before and during operation. Since new faults are usually introduced to the code during updates, a reliable and verified software update should also be guaranteed (R1.3). In addition, the interaction between software and hardware should be robust enough to guarantee safe operation (R1.4).

Some external supporting systems might be involved during the operation of AMS, e.g., a UUV requires an underwater navigation system, and MASS and UUV may require a control basis/center for remote supervision and control. Therefore, if any external supporting system is involved, the communication and interaction between them and the AMS should be adequate and reliable (R1.5). The interaction and communication with other marine stakeholders and environments, such as other ships, marine structures, and the Vessel Traffic Service (VTS), should also be considered during operation (R1.6).

Autonomous ships may switch between various operational modes with different Levels of Autonomy (LoA) due to the rapidly changing environment or complex nature of tasks (Thieme et al., 2018; Yang et al., 2020b). Other AMS, such as ROV, may also need to operate in an adaptive autonomy/mode (Hegde et al., 2018b; Yang et al., 2020b). Reliable and adequate provisions for adaptive autonomy/mode are required in the AMS operation (R1.7). Human-machine interactions and cooperation are expressed by various LoAs, and each level specifies a different degree of operation between fully manual operation and highly autonomous operation (Vagia et al., 2016). Although the ultimate goal is to have highly autonomous systems, human operators are still required for each AMS, currently and in the near future. Thus, it is necessary to have an accessible and affordable human–machine interface (R1.8).

Table 2 summarizes the requirements for a general online risk model. The risk spectrum of the system or operation is expected to be measured by utilizing various sources of data, including historical data, expert knowledge, and especially the monitoring data from sensors (R2.1). With the help of online data, online risk models are expected to provide a real-time risk picture and pre-warnings of possible operational deviations (R2.2). By capturing data and information during the operation

Table 2
Requirements for a general online risk model.

Requirements	Description
R2.1	Utilize various sources of data, especially the monitoring data from sensors, in order to provide the risk spectrum of the system or operation
R2.2	Dynamic in order to capture the quick changes in operation
R2.3	Update models with new information, data, and scenario for better risk evaluation and emerging risk
R2.4	Capture the possible changes of involved subsystems or components and their impacts on risks during the operation
R2.5	Efficiently identify RIFs that need to be monitored online or in real time during operation
R2.6	Effectively model the correlation among identified RIFs to estimate the overall risk level
R2.7	Capture the uncertainty in the model, especially the uncertainty caused by sensors and the data fusion algorithm

using the monitoring technique, an online risk model should be able to update the model, in terms of both the model itself and the type of input data, for better risk evaluation and emerging risks (R2.3).

The system or operation may involve different subsystems or components in different phases in a task. The relevant data that needs to be considered and monitored in the risk model may change over time. In addition, due to changes in the interaction between the subsystems, new hazards may evolve, and the acceptable risk level of the operation may also change accordingly. An online risk model should be able to reflect these changes in different phases during an operation (R2.4). A risk model needs to identify factors that may affect the level of risk. These factors are called Risk Influencing Factors (RIFs), which are defined as "a set of conditions which influence the level of specified risks related to a given activity or system" (Rosness, 1998). By monitoring the states of the RIFs, early warnings about possible deviations from the normal operating envelope of a system can be provided (Utne et al., 2020). In order to efficiently monitor the system and provide an accurate risk evaluation, an online risk model should efficiently identify RIFs that need to be monitored online or in real time during operation (R2.5). The last two requirements are similar to those of traditional risk models and the existing dynamic risk models. The online risk models should be able to effectively model the correlation among identified RIFs and reflect the overall risk level of the system (R2.6). The uncertainty should be properly handled in online risk models, especially the uncertainty caused by sensors and the data fusion algorithm that is caused by the increase in the use of monitoring techniques (R2.7).

3.2. Evaluation criteria for online risk modeling of AMS

The evaluation criteria for online risk models of AMS are derived based on the requirements identified in Tables 1 and 2. The criteria are developed considering that the online risk model should be used for the AMS itself to operate autonomously, and/or for the human operators to monitor the operational situation. Table 3 summarizes the criteria identified for the online risk model of AMS. The evaluation criteria reflect the aspects that online risk models of AMS need to cover. The current list of evaluation criteria can be used to assess the validity and effectiveness of online risk models. It is also expected to be used as a guide to check whether any important aspects of the online risk model are missing and what new information should be included.

Table 3
Evaluation criteria for online risk modeling of AMS.

Identifier	Criteria for online risk modeling of AMS	Addressed requirements
C1	Inclusion of maintenance and reliability aspects of system performance	R1.1, R1.2
C2	Inclusion of the performance of software and control algorithm	R1.1, R1.3
C3	Inclusion of the performance of the interaction between software and hardware	R1.1, R1.4
C4	Inclusion of the performance of the interaction between AMS and external supporting system	R1.1, R1.5
C5	Inclusion of the performance of the communication between AMS and environment	R1.1, R1.6
C6	Inclusion of the hazards and possible changes in risk models caused by adaptive autonomy/mode or the change of involved subsystems	R1.1, R1.7, R2.4
C7	Inclusion of human–machine interaction	R1.1, R1.8
C8	Inclusion of security issues	R1.1-R1.8
C9	Inclusion of various sources of data to estimate the risk level, especially sensor data	R1.1, R2.1
C10	Level of knowledge (in both the studied system and risk) needed for analysis	R2.1
C11	Be able to update risk level with new information/ data	R2.2
C12	Be able to deal with emerging risk (the way that the model is changed and/or updated with new data)	R2.3
C13	Be able to efficiently identify RIFs that need to be monitored online or in real time during operation	R2.5
C14	Be able to effectively model the correlation among identified RIFs	R2.6
C15	Be able to deal with the uncertainty, especially the uncertainty from the sensor and real-time data or the data fusion algorithm	R2.7

4. Development of online risk models

The first step of risk analysis is to identify what can go wrong, and relevant methods include Hazard Identification (HAZID) and the STPA. The PHA is an extended version of HAZID that also addresses the likelihood and consequences, usually in a semi-quantitative manner. Procedural HAZOP is used to review procedures and operational sequences. Hence, the STPA, PHA and Procedural HAZOP may therefore provide a desirable foundation for developing an online risk model.

Some previous studies have been conducted on the comparison of different methods, such as STPA and HAZOP (Sultana et al., 2019) or STPA and FMEA (Rokseth et al., 2017), against various aspects to demonstrate how one method can be used to replace another, or how one method can be used as complementary to another one. The current study, however, aims at analyzing the applicability of different methods to the online risk modeling of AMS, identifying advantages and disadvantages over the identified criteria and determining appropriate methods based on the analysis result. Information from each method that can be utilized to further develop an online risk model is also identified.

The PHA is usually used to identify hazards and potential accidents in the early stages of system design and has been successfully applied to safety analysis in many fields, such as process plants and offshore marine systems (Rausand, 2013; Vinnem and Røed, 2019). The term "preliminary" reflects that the analysis results are usually refined through additional and more thorough studies when more information on the system becomes available. Hence, a PHA is typically used to provide an initial risk picture for the system, but may also be used as a stand-alone analysis. Still, when a more comprehensive risk assessment is necessary, the analysis results can also be used to screen events for further research, making it possible for them to become the basis of online risk models.

A HAZOP study is a structured and systematic hazard identification process that examines how a system may deviate from the design intent and results in hazards and operability problems that may represent risks to personnel or equipment. The studied system is divided into several simpler sections called "study nodes" that are analyzed one by one later (Rausand, 2013), by using a set of guidewords and process parameters. The analysis is carried out by a group of experts from different research areas (a HAZOP team) in a series of brainstorming sessions. The HAZOP approach was initially developed to be used during the design phase, but can also be applied to systems in operation. Several variants of the original HAZOP approach have been developed (Rausand, 2013). Procedural HAZOP is considered a powerful tool for risk assessment of new or changed operations and is applicable for all activities where an operational procedure is used (Vinnem and Røed, 2019).

STPA is a hazard analysis method mainly based on the idea of System-Theoretic Accident Model and Processes (STAMP), in which safety is controlled by enforcing constraints on the system behavior (Leveson, 2011). Unsafe interaction among the components in a system is believed to be the main reason leading to an accident, instead of considering that the accident is the result of a chain of component or event failures. The hazardous events occur due to the absence, presence, or the improper timing of control actions. The method is usually selected due to its ability to model complex interactions. In general, the process of STPA consists of the following steps:

- Step 1: Define the purpose of the analysis, including system to be analyzed and also the analysis boundary. Hazardous events at system level and safety constraints need to be identified as well.

- Step 2: Develop the hierarchical control structure of the system to be analyzed. Interactions among the components are represented by control actions and feedbacks.
- Step 3: Identify Unsafe Control Actions (UCA) that violate the safety constraints.
- Step 4: Develop loss scenarios in which UCA may occur and identify their causes.

The methods mentioned above cannot be used for developing online risk models directly, but based on such methods, more detailed risk modeling can be performed, using, for example, BBN, FTA, ETA, Hybrid Causal Logic (HCL), simulation-based approaches, etc. The current study aims to analyze the applicability of PHA, STPA and Procedural HAZOP as the starting point for online risk modeling of AMS, but the development of a comprehensive model is outside the scope of the current work.

5. Case study

5.1. Under-ice operation of AUV

An AUV under-ice operation is used as a case study to investigate how PHA, Procedural HAZOP, and STPA contribute to fulfilling the criteria for an online risk model of AMS.

As a part of the Nansen Legacy project (The Nansen Legacy), AUVs are used to collect environmental data of the oceans, and under the ice in the Arctic region in the near future. Arctic operations, however, involve risks related to loss of the vehicle and mission abortion, due to the harsh environmental conditions for vehicles and human operators and difficulties in AUV navigation. Loss of the vehicle and abrupted or aborted missions are costly due to the high expenses related to the vessels used for the field cruises, but the consequences are also related to the failure to collect the data used for ocean monitoring and science. Furthermore, loss of the vehicle has a negative environmental impact in terms of adding to the “garbage” in the oceans. Compared to the traditional AUV operation, the difficulties with under-ice operations include but are not limited to the following:

- Logistical challenges due to remote areas and limited infrastructure
- Harsh environmental conditions for operation, such as the low temperature and the presence of ice
- Navigation challenges of the Arctic area
  - o The large vertical component of the magnetic field reduces the accuracy of the magnetic compass
  - o The low horizontal component of the Earth’s rotation reduces the accuracy of the gyroscopic compass

To improve the safety and robustness of under-ice operations with AUVs in the Arctic, an online risk model is needed to provide decision support for the human operators and the AUV itself. In this research, the NTNU REMUS 100 AUV has been selected to perform under-ice operations, considering its robustness and previous under-ice track record. Details about this AUV can be found in (Norgren et al., 2020). Considering that most AUVs have similar characteristics to the NTNU REMUS 100 AUV, the results obtained from this work should also be valid for most AUVs.

5.2. System description

Fig. 1 demonstrates the schematic diagram of the preliminary design of the AUV operation using under-ice navigation buoys, currently under development at NTNU. The design aims to deploy navigation buoys along the planned AUV transect, providing AUV navigational support during the mission. The concept of Single Transponder Navigation (STN) is applied, making use of the short-term positioning accuracy provided by the high-performance dead-reckoning navigation system in the AUV and bounded long-term accuracy provided by the buoys system (Norgren et al., 2020). When the AUV is operating under ice, it will measure the distance to the buoys via acoustics. In addition, since the buoys may drift with ice and ocean currents, the buoys need to obtain the position through the Global Navigation Satellite System (GNSS) and transmit it to the AUV via acoustics. Combined with the position estimation from the high-performance dead-reckoning navigation system in the AUV, relatively good navigation performance can be obtained. Equipped with several environmental sensors, the AUV is expected to collect data. More detailed information on the system can be found in the study by Norgren et al. (2020). With the help of the designed system, the AUV operation in the current case study aims to search for the temperature gradient and follow the route that decreases with the temperature gradient. PHA, Procedural HAZOP, and STPA have been applied to analyze the possible risks of the operation.

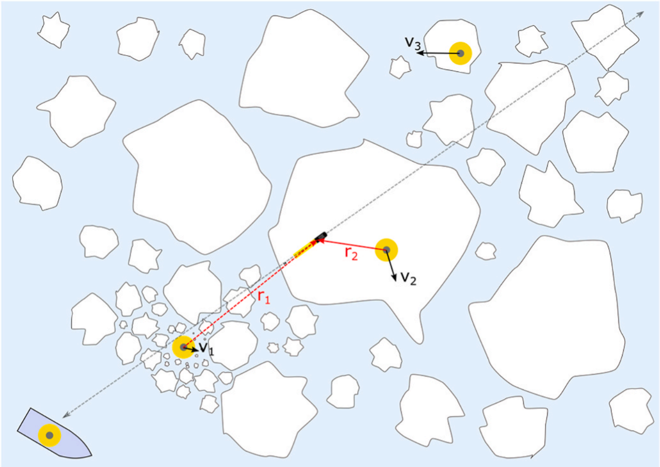


Fig. 1. AUV operation using the under-ice navigation buoys, designed by Norgren et al. (2020).

6. Main results and findings of the case study

The current section presents the main results and findings from the three methods.[1] The analyses involve risk analysts and AUV experts with experiences from several previous Arctic AUV operations with vehicles from NTNU, especially from a research cruise to the North Barents Sea in November 2019. The AUV operations provide the participants in the analyses with field experience on environmental conditions and the challenges of AUV operation in the Arctic, including underwater navigation challenges, technical and operational failures, and logistical challenges.

6.1. Main findings from PHA

The PHA was conducted through three PHA workshops, which gathered people from different fields of expertise, i.e., risk assessment and AUV operation. The workshops resulted in several hazardous events, which were identified and analyzed with respect to their assumed frequencies and expected consequences. In the analysis, Table 4 and Table 5 were used for the categories of frequency and consequence, respectively. The expected consequences were identified considering the principle of the credible worst-case before any risk reduction measures have been implemented. Fig. 2 presents the risk matrix used in the current risk analysis, where the risk index is a semi-quantitative measurement of risk and defined as “the logarithm of the risk associated with the event and is found by adding the frequency class

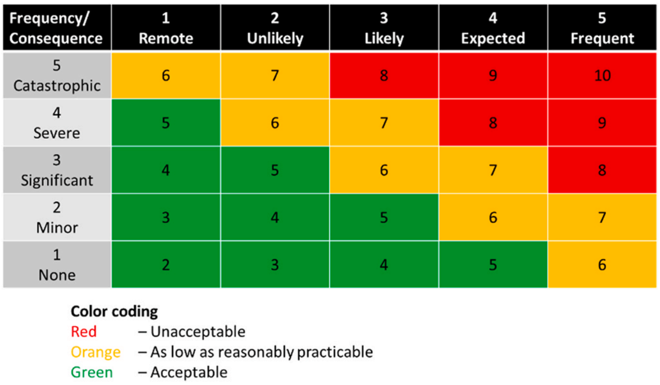
[1] More detailed results of the analyses can be provided by contacting the corresponding author (Ruochen Yang. ruochen.yang@ntnu.no)

Table 4
Frequency categories for use in the current PHA.

Index	Category	Frequency (per operation)	Description
5	Frequent	>1	The event is likely to occur more than once per operation.
4	Expected	Around 1	The event may occur once per operation
3	Likely	1–0.1	The event may occur once per operation/ten operations
2	Unlikely	0.01−0.1	The event will be most likely not to occur
1	Remote	<0.01	The event is unlikely to occur

Table 5
Consequence categories for use in the current PHA.

Index	Category	Consequence	Description
5	Catastrophic	Loss of AUV/injuries to the operators	Loss of time (over 100 days), over 1,000,000 NOK
4	Severe	Major damage to the system (AUV/buoys)/loss of several buoys	Loss of time (100 days), 500,000 NOK-1,000,000 NOK
3	Significant	Mission failure (unable to repeat)/no data/loss of one buoy	Loss of time (ten days)/data, 500,000 NOK
2	Minor	Minor influence of mission/ unacceptable data/Minor damage to the system	Loss of time (one day)/ data
1	None	No damage/influence	No loss of time/data



Frequency/ Consequence	1 Remote	2 Unlikely	3 Likely	4 Expected	5 Frequent
5 Catastrophic	6	7	8	9	10
4 Severe	5	6	7	8	9
3 Significant	4	5	6	7	8
2 Minor	3	4	5	6	7
1 None	2	3	4	5	6

Color coding
Red – Unacceptable
Orange – As low as reasonably practicable
Green – Acceptable

Fig. 2. Risk matrix for use in the current PHA.

of the event with the severity class of the event" (Rausand, 2013). Different colors represent the level of acceptance.

The PHA focused on the hazardous events related to the REMUS AUV system's technical hazards, technical hazards of the navigation buoys system, environment, traffic and operational hazards, and human error. Table 6 summarizes the most hazardous events and their possible causes, consequences, and suggested risk reduction measures. It is found that the technical hazards of the REMUS AUV and the harsh environment contribute most to the risks of the AUV's under-ice operation.

In terms of the technical hazards of the REMUS AUV, the AUV's navigation and communication system module failure and software failures are considered to be the most hazardous events, which may directly lead to a loss of the AUV. Risk-reducing measures are proposed to mitigate these risks. Considering the rapidly change of operating environment, unexpected navigation challenges in the Arctic, and relatively little experience with under-ice operation, the navigation and communication system module and other components should be fully tested under different operating conditions before operation. In order to avoid unwanted software failure, software verification and testing, especially for own-developed software (control algorithm and software configuration), should be carried out before operation.

Unacceptable hazardous events associated with the environment include the low maneuverability caused by the strong current and the potential accidents caused by Arctic sea ice, such as the collision with ice or stuck under ice. In order to mitigate the risk, enough preparations are needed to retrieve and salvage the AUV, such as the acoustic pinger for pinpointing AUV's location and tools for cutting ice. These also require human operators to be well trained and familiar with the retrieve and salvage process in the Arctic.

6.2. *Main findings from procedural HAZOP*

The entire AUV operation is divided into five main phases in the current study, including pre-deployment, deployment, operation, recovery, and post-deployment. It was summarized by the designer of the system, who has over seven years of work experience in AUV operation. Procedural HAZOP was applied to identify deviations from the way the system is intended to function: their causes, and all the hazards and operability problems associated with these deviations. Each main step in the operational procedure is regarded as a "study node" in the current HAZOP work. A list of guidewords used for identifying deviation was agreed on by all experts before analysis, as shown in Table 7.

The current Procedural HAZOP analysis was conducted through three HAZOP workshops, gathering same analysts as the PHA workshops. Table 8 shows examples of the analysis results. Though most hazardous events related to human operators and operational procedure are identified in the phases of pre-deployment, deployment, and recovery due to the operators' high involvement, these hazardous events may affect other phases or even the whole operation phases as well, and lead to an unacceptable consequence. For example, the failure of testing of communication between buoys may result in a non-functional buoy and then cause navigation failure during the operation phase and lead to the loss of AUV. The results in Procedural HAZOP highlight the importance of the proper testing and verification of software and hardware in AUV and buoys and adequate preparation for environmental and operational challenges before the operation phase.

6.3. *Main findings from STPA*

The STPA analysis was initially performed by the first author and then reviewed and revised by the same analysts as in PHA and Procedural HAZOP workshops. The STAMP Workbench software (Information-technology Promotion Agency, 2021) was used to develop the control structure and further analysis.

A hierarchical control structure diagram for AUV under-ice operation is shown in Fig. 3. It demonstrates the main interactions between each component in the system. Available control actions from the controllers are represented by the red arrows, while feedback signals provided by actuators are represented by the blue arrows. Generally, the operation may involve the human operators, the AUV, the navigation system, the supporting system, etc. In this study, the AUV, operators, and navigation buoys systems are selected as the main elements in the system.

Given the hierarchical control structure developed in Fig. 3, the UCAs can be determined by considering all the control actions. For each control action, four categories of UCAs are considered to identify the actions that have the potential to cause hazards (Leveson, 2011). Table 9 presents an excerpt of the UCAs found in the case study. No UCAs that were caused by applying a control action for too long or for stopping too early were identified.

The causal scenarios can be identified by analyzing how the identified UCAs may occur. The current case study takes UCA5-P-1 (Navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission) as an example to show the

Table 6
Unacceptable hazardous events and risk reduction measures obtained from PHA.

Hazardous event	Cause	Consequence	Risk			Risk-reducing measures/ actions
			Fr.	Cons.	RI	
Technical hazards of REMUS AUV system						
Navigation and communication system						
Bad Long Baseline (LBL) range measurement: Multipath etc.	Multipath from ice	Poor navigation signal or loss of navigation, making AUV fail to identify its position, follow the desired path, and collect satisfactory data	5	3	8	
Software system						
Failure of software system on operator's computer	Low battery on computer	Unable to connect AUV, probably leading to loss of AUV	3	5	8	Conduct testing in different weather conditions
Failure of software system on AUV	Bad configuration of software (modified by operators)	Unable to provide desired control/ navigation, probably leading to loss of AUV	3	5	8	Software verification, especially testing own software before operation
Failure of self-designed software (inside AUV)	Bad configuration of software	Unable to provide desired control/ navigation, probably leading to loss of AUV	3	5	8	Testing of configuration before operation
Environmental hazards						
AUV is unable to reach the planned retrieval point	Strong water current causes AUV to fail to reach /remain in the planned retrieval point.	Failure/hard to retrieve AUV, leading to loss of operation time or AUV	3	5	8	Testing and practice/training of the operators. Prepare pinger for triangulation pinpoint; prepare chainsaw for fast hole. Use snowmobile. Know direction of current.
AUV gets stuck under ice	AUV battery depleted	Unable to float to surface freely, probably loss of AUV due to difficulties in pinpointing position	3	5	8	Testing and practice/training of the operators. Prepare pinger for triangulation pinpoint; prepare chainsaw for fast hole. Use snowmobile. Know direction of current.
Collision/contact with floating ice	Ice in diving/surfacing area	Damage to antenna, making it difficult to connect to operators, probably leading to loss of AUV	3	5	8	Test and verify the reliability of antenna before operation
Human error						
Insufficient test before operation	Human error	Faults are not recognized during planning phase. Emerging risk may lead to loss of AUV	4	5	9	Set up a test checklist for operators, and make sure that operators follow the testing procedure before operation

derived loss scenarios and causal factors, as presented in Table 10. Five possible scenarios are identified for UCA5-P-1, taking into account unsafe controller behavior (such as inadequate control algorithm), inadequate feedback and information, etc. Since AUV relies on navigation buoy for navigation, AUV may provide unacceptable inaccurate states estimation during the mission if the inaccurate position is provided to AUV, or the position of buoys is not updated to AUV. In addition, the failure of software, navigation system module in AUV, or the failure of measuring accurate depth, altitude and AUV speed might also lead to the unacceptable inaccurate estimated position and heading. Given the loss scenarios identified, a more detailed analysis can be performed to identify the casual factors. Related casual factors include the distance between buoy and AUV, reliability and uncertainty of acoustics of navigation buoy and AUV, reliability and uncertainty of GPS signal, etc.

7. Discussion

7.1. *Main risks and implications for improved engineering design, operational procedures and further research*

The results from the above analyses may assist designers and operators to improve the safety and robustness of vehicles and operations in the Arctic in the future.

Firstly, a relatively high number of potential hazardous events and/ or UCAs can be traced back to the failure of the physical components in the AUV or buoys. In conventional AUV operation in open water, a fail-to-safe mechanism of floating to the surface is common when any fault is detected, such as a leakage in the AUV. Instead, due to the possible existence of ice coverage in AUV under-ice operation, a commonly used fail-to-safe mechanism is to park the vehicle on the bottom and wait until it is guided to a safe location (Ferguson, 2008). However, when any critical component fails during the operation, such as the leakage in the

Table 7
Guidewords used in current Procedural HAZOP study, based on (Broadleaf, 2018; IEC, 2016).

Guidewords	Topics for discussion in the workshop
No action	Step is missed or omitted; intended AUV operation did not occur; action impossible; AUV or supporting system (buoy or research vessel) not ready
Less action	Human operator does less than intended; hardware does not perform as required; not enough time to complete the step
Wrong action	Human operator does the wrong thing, starts the wrong job, reads the wrong instructions; personnel perform different or out of date procedure; perform two or more steps at the same time
Out of sequence	Human operator misses out a step; carries out a step before it should occur, or after it
More time	Human operator takes longer than necessary over action (leaves something running and gets distracted); starts next action later than expected
Less time	Human operator carries out action too quickly; starts next action earlier than expected
No information	No information or feedback from the process or operation; procedure does not specify expected performance; no specified actions for emergencies
Wrong information	Information provided is wrong, out of date or contradictory (oral instruction vs. written, other procedures or steps within this procedure)
Clarity	Step is confusing; words are confusing; readability; poor procedure form layout; written in non-English language; not clearly understandable
Training	Adequate training; level of certification required and provided for this step; procedure control (issuing, updating, revisions, overriding, communication, distribution, and acknowledgment, retraining)
Abnormal conditions	Emergencies; recovery from abnormal situations; utility failure; severe or unusual weather; deviation from procedure; make-shift operations
Safety	Personnel protection; Occupational Health and Safety (OH&S) law compliance; industrial hygiene issues; environmental considerations; fire, explosion or chemical release potential

AUV or the physical failure of the propeller, waiting on the bottom for a period to find the safe location might be challenging.

The predefined fail-to-safe mechanism may not be performed as intended, and this may directly lead to the loss of AUV considering the difficulties to salvage under-ice AUV. Therefore, compared to the operation in open water, more severe consequences can be incurred if there is any failure of the physical components in the AUV or buoys. From the perspective of engineering design, the operation of AUVs in the Arctic requires more reliable and robust physical components. According to the analyses results, adequate testing and verification of the components' reliability in various operating environment before operation are suggested. Also, a more effective fail-to-safe mechanism is helpful to deal with the challenge of retrieving the AUV.

Compared with hardware failure, operators may be more interested in the risk related to software or control algorithm in the operation of AMS since these contribute most uncertainty and unexpected hazardous events. A good example is the challenges of underwater navigation in the Arctic. All three methods identify navigation failure or error as a major issue. Several underwater navigation difficulties may pose challenges related to this, for example, the multiple paths from ice and seabed due to successive reflections at the interfaces when signals transmit, high ambient acoustic noise caused by either natural or man-made sources, and lost signal caused by buoys drift out of the acoustic range.

An inadequate algorithm for calculating and mitigating the navigation uncertainty can result in the loss of AUV, since it may be difficult for the AUV to determine an accurate location for retrieval. Therefore, a more robust algorithm is needed to deal with the navigational uncertainty. Testing and verification of all onboard software should be performed to ensure quality of software application and design. In terms of drifting buoys, more reliable algorithm can be applied to simulate and predict the drifting of the buoys to prevent the buoys from drifting out of the acoustic range during operation. Deploying the buoys in a relatively closer distance to possible AUV operating path can also be effective. In addition, several RIFs or risk indicators related to underwater navigation might be crucial for limiting the uncertainty in navigation, such as the distance between AUV and buoy and standard deviation of reported buoy's position, an onboard online risk model that can capture these values can be helpful to reduce the risk of operation.

Although the human operators are not directly involved and have little control of the vehicle during the operation phase of AUV mission, hazards in operational procedures will still have a great importance to the safety of AMS operation. According to the results, these can be associated with inadequate system design, defective software development, insufficient preparation and testing, improper operation steps and behaviors, limited work schedule, etc. The three analyses highlight the importance of adequate preparation for environmental and operational challenges. A packing list of necessary equipment for operation and recovery and a checklist operational procedure should be provided to human operators to avoid missing of necessary equipment and operational steps. Due to the logistic challenges of the operation in the Arctic, such as limited time of operation and recovery caused by unexpected challenges with testing and deploying AUV or buoys in the Arctic, schedule change of research vessel, etc., a good communication with crew of research vessel and cruise leader should be ensured and a possible backup plan for operation is necessary.

The above indicates a need for further research in the domain of autonomous operation in the Arctic. Engineering design and operational procedures, for example, need to be improved. Since the scope of the paper is not on the design of the AUV, the next subsections focus on the use of the analyses results for online risk modeling only.

7.2. Applicability of using the results in online risk modeling of AMS

This section analyzes the applicability of the three methods to the identified criteria of online risk modeling of AMS. Table 11 summarizes the main findings from the analysis results, and the following subsections present detailed arguments and observations supporting these assessments based on the analysis from the case study. The current study does not rank the importance of these criteria, since each derived criterion covers important aspects of online risk models. However, stakeholders may be more interested in criteria that reflect the main difference between conventional marine systems and AMS, or between traditional risk models and the online risk model (for example, criteria C2, C3, C4, C6, C8, C9, C11, C15) than other criteria. Analysts may focus on different criteria when developing an online risk model, depending on the type of AMS, available data, etc.

Generally, compared to the other two methods, STPA shows better applicability in terms of the number of criteria fulfilled. The STPA results demonstrate a more detailed analysis of the risk caused by the software and control algorithms due to its ability to handle the risk caused by unsafe interaction. The visualization of the interaction between the AMS and external systems is very valuable in the analysis of AMS operation, in which these interactions might bring more issues compared to the operation of traditional marine systems. Although the current study does not consider the security issue and adaptive mode, other studies show the method's capability. The main disadvantage of STPA for an online risk model is that it provides a list of hazardous events without any ranking or quantification of the risk, making it difficult for analysts to determine which RIFs should be selected for inclusion in the model.

Table 8
Examples of hazardous events and risk reduction measures from Procedural HAZOP.

Guide word	Deviation	Possible causes	Consequence	Existing control	Action required
Pre-deployment					
Less action	Operators do not/unable to test acoustic communication between all buoys and AUV (successfully)	Test repeated in the same buoy	Navigation failure during operation/ mission abortion/loss of vehicle/unable to send command/limit the navigation range	Checklist for testing before operation (make sure operator remembers and follows the checklist)	Mark each buoy with label/bring the ranger and test
Deployment					
No action	Failure to deploy slave-buoys at desired locations	Ice condition/no (not enough) required tools/	Failure to provide navigation to AUV/ no deployment of AUV operation/ mission delay/drift of the buoy out of range	Send list of required tools before mission; check the required tools on board/check availability/ check expected ice thickness information with other groups	Bring enough tools for ice on the cruise
Less action	Operators do not/unable to test communication (Very High Frequency (VHF) radio, iridium) between all buoys successfully	Hardware failure (e.g. transmitter, power for drive board)/ software failure/forget/configuration issue/may not be successful if not tested in water	Delay mission start/unable to operate/ navigation failure during operation	Checklist (make sure operator remembers)/ avoid operation if there is issue	
More time	Take longer time to deploy buoys and AUV	Ice condition; weather; polar bear; not properly prepared; not good communication with other research teams (the mission is put on the waiting list for longer time)	Mission delay, less operation/ investigation time	Be polar bear guard ourselves; pass the shooting course; good communication with other research teams	
Operation					
More time	Delayed report of the AUV status	Acoustic multipath; not good acoustic communication condition; condition of the buoy software	Normally not critical; don't understand the information and make wrong decision (battery level is reported late)	Get Conductivity, Temperature, Depth (CTD) profile to know the acoustic communication condition	
Wrong	information	Status of AUV is not correctly reported during the mission (critical error is detected when it is normal)	Software error; sensors issue	Mission abort; lose data; lose AUV	Proper test/ preparation; need to recover AUV if the reported status is not feasible
Recovery					
No action	Operators unable to recover the AUV/ buoys	Bad handling; ice condition; buoy is frozen; ship doesn't allow its recovery (bad weather, polar bear); navigation problem may cause AUV to be unable to pinpoint the place to recover (navigation error should be fine)	Longer to recover; damage to the AUV/ buoy; deplete AUV battery (and possibly lose AUV due to no communication then)	Good handling; be aware of the ice/bear condition, weather condition; bring enough equipment for recovery (rope, tripod)	

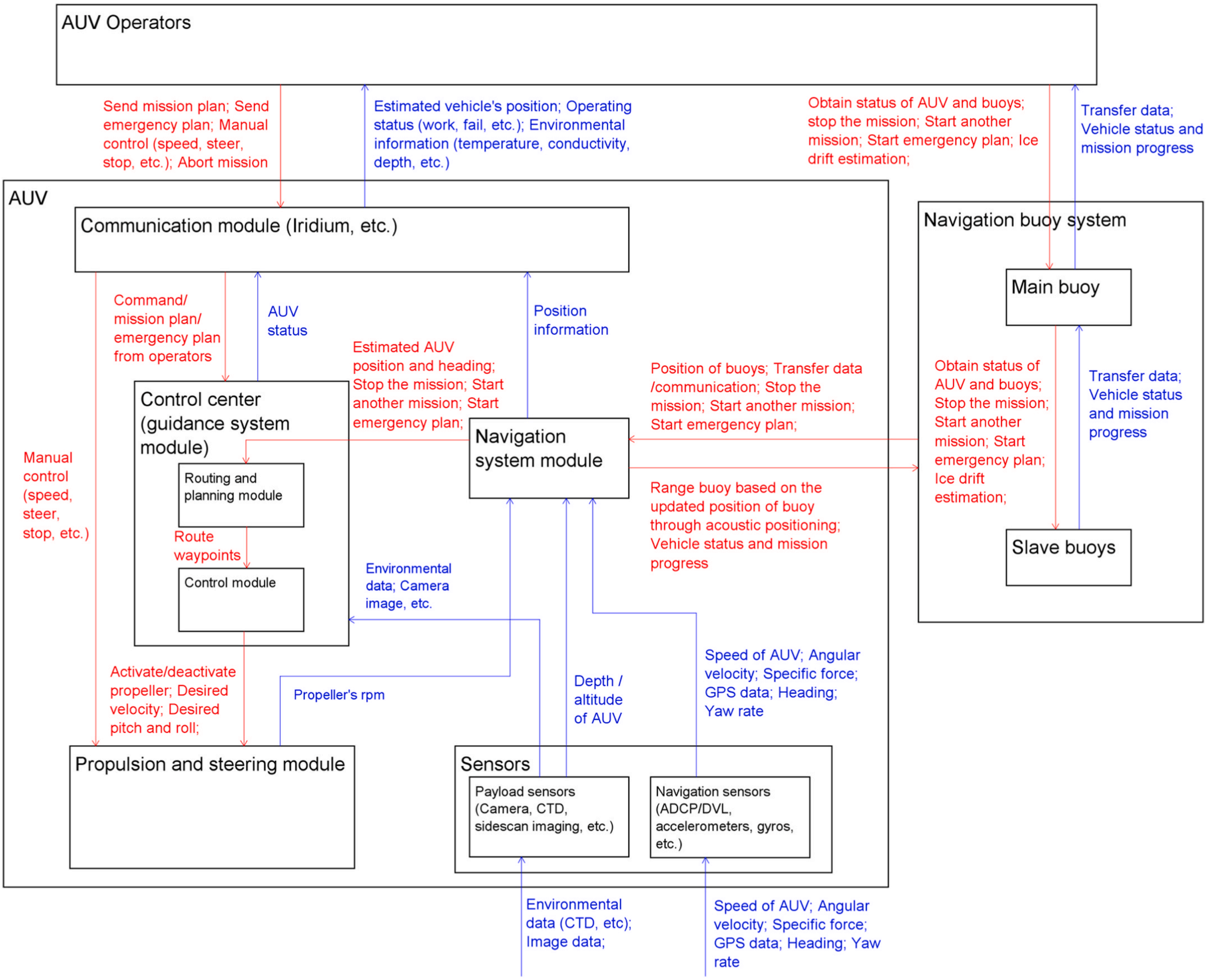

AUV Operators
Send mission plan; Send emergency plan; Manual control (speed, steer, stop, etc.); Abort mission
Estimated vehicle's position; Operating status (work, fail, etc.); Environmental information (temperature, conductivity, depth, etc.)
Obtain status of AUV and buoys; stop the mission; Start another mission; Start emergency plan; Ice drift estimation;
Transfer data; Vehicle status and mission progress
AUV
Communication module (Iridium, etc.)
Command/ mission plan/ emergency plan from operators
AUV status
Position information
Estimated AUV position and heading; Stop the mission; Start another mission; Start emergency plan;
Control center (guidance system module)
Routing and planning module
Route waypoints
Control module
Manual control (speed, steer, stop, etc.)
Navigation system module
Position of buoys; Transfer data /communication; Stop the mission; Start another mission; Start emergency plan;
Range buoy based on the updated position of buoy through acoustic positioning; Vehicle status and mission progress
Navigation buoy system
Main buoy
Obtain status of AUV and buoys; Stop the mission; Start another mission; Start emergency plan; Ice drift estimation;
Transfer data; Vehicle status and mission progress
Slave buoys
Environmental data; Camera image, etc.
Activate/deactivate propeller; Desired velocity; Desired pitch and roll;
Propeller's rpm
Depth / altitude of AUV
Speed of AUV; Angular velocity; Specific force; GPS data; Heading; Yaw rate
Propulsion and steering module
Sensors
Payload sensors (Camera, CTD, sidescan imaging, etc.)
Navigation sensors (ADCP/DVL, accelerometers, gyros, etc.)
Environmental data (CTD, etc); Image data;
Speed of AUV; Angular velocity; Specific force; GPS data; Heading; Yaw rate

Fig. 3. Hierarchical control structure diagram for AUV under-ice operation.

The PHA results provide a good understanding of the system with the relatively low knowledge and experience level required. It performs well with respect to identifying reliability and maintenance aspects, but it may be challenging to provide a detailed analysis of software-related failures without a detailed hazard list. In terms of the environmental influence, PHA provides a clearer view than STPA by explicitly considering the impact of various environmental factors, making it a preferred method for operations in harsh environments like the Arctic or space. Semi-quantitative results enable analysts to rank the importance of identified hazardous events and make PHA easier to use as a basis for further building of online risk models. The ability to handle adaptive autonomy, software-related failures, and security issues makes it difficult to become an ideal online risk modeling method. However, considering the acceptable results obtained and less time spent on PHA, it could serve as a basis for developing STPA, where developing the control structure and UCAs is challenging for inexperienced analysts. The analysis results can help analysts better understand the system and its interaction with other systems, thereby developing a more satisfactory control structure and UCAs.

In terms of evaluation criteria, the behavior of Procedural HAZOP is unsatisfactory in aspects such as the ability to deal with the reliability-related issue, software-related issues, adaptive autonomy, and security issues. However, it does provide some results that are not covered by the other two methods. With a detailed operational procedure, Procedural HAZOP mainly focuses on the behavior of human operators. In the analysis of the environmental impact and the human-machine interaction, it provides much better results related to operators' behavior than PHA or STPA, which makes it an excellent complementary tool for them.

According to the analysis results in Table 11, none of the three methods contribute to all the online risk model criteria. More details are provided in the next subsections.

7.2.1. Inclusion of maintenance and reliability aspects of system performance

In general, both PHA and STPA show good coverage in terms of the aspects of reliability and maintenance.

In PHA, this aspect is mainly reflected in technical hazards related to both the REMUS AUV and the navigation buoys system. The system is broken down according to its physical structure during the analysis. For example, in the navigation and communication system of the REMUS AUV, various failure modes of the GNSS system, Long Baseline (LBL) transducer, Inertial Navigation System (INS), Acoustic Doppler Current Profiler (ADCP), etc., can be well listed and analyzed. Based on the experience from the current study, operators or AUV experts are familiar

Table 9
Examples of UCAs identified for the AUV under-ice operation in the Arctic.

CAs	Not providing	Providing causes hazard	Too early/Too late
Estimated AUV position and heading [from navigation system module to routing and planning module]	(UCA5-N-1) Navigation system module does not provide estimated position and heading during the mission [SC1][SC2][SC5][SC6] (UCA5-N-2) Navigation system module does not provide command to stop the mission/start another mission when the current mission is finished [SC5][SC6] (UCA5-N-3) Navigation system module does not provide command to start emergency plan when operators believe there is an issue and send the command to start emergency [SC1][SC2][SC3]	(UCA5-P-1) Navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission [SC1][SC2][SC5][SC6] (UCA5-P-2) Navigation system module sends the command to stop the mission/start another mission/emergency plan when the current mission is working smoothly and successfully [SC5][SC6]	(UCA5-T-1) Navigation system module sends the command to stop the mission/start another mission too late when the current mission is already done [SC6] (UCA5-T-2) Navigation system module sends the command for emergency plan too late when the failures/mistakes are detected [SC1][SC3]
Activate/deactivate propeller; desired revolutions per minute (rpm) for each thruster; desired pitch and roll; perform emergency plan [from control module to propulsion and steering module]	(UCA6-N-1) Control module does not provide desired rpm (higher or lower) when the vehicle is close to other objects [SC1][SC2][SC6] (UCA6-N-2) Control module does not provide desired pitch/roll (higher or lower value) when the vehicle should follow the designed path [SC5][SC6] (UCA6-N-3) Control module does not provide "performing emergency plan" command when failure/pre-defined situation occurs [SC1][SC3]	(UCA6-P-1) Control module provides "Mission aborts, and emergency plan" command when the AUV is working smoothly and successfully [SC6] (UCA6-P-2) Control module activates propeller when the vehicle is already in the designed location and should stop for a while for next stage [SC5][SC6] (UCA6-P-3) Control module activates propeller when the vehicle stops and is close to other objects [SC1][SC2] (UCA6-P-4) Control module deactivates propeller when the vehicle is on the way to the designed location [SC5][SC6] (UCA6-P-5) Control module provides undesired pitch/roll (higher or lower value) when the vehicle should follow the designated path (temperature gradient) [SC5][SC6]	(UCA6-T-1) Control module deactivates propeller too late when the vehicle is close to other items [SC1][SC2]

with the physical components of the system; thus, almost all the physical components can be included with the help of design details and expert experience. With a typical checklist of possible hazards, PHA can provide a complete and detailed analysis in terms of reliability and maintenance.

The control structure used in STPA is a functional model, not a physical model like a physical block diagram, and the control actions or feedbacks do not necessarily reflect the physical interactions (Leveson and Thomas, 2018). Given that the UCAs identified are based on the control structure, the aspects of reliability and maintenance can be considered when analyzing the loss scenarios for UCAs. This process can be done by asking: 1) why would UCAs occur; and 2) why would control actions be improperly executed or not executed, leading to hazards. For example, the navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission (UCA5-P-1) if the correct depth or altitude of the AUV is not provided, which can be

Table 10
Loss scenarios for UCA5-P-1: Navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission [SC1][SC2][SC5][SC6].

No.	Causal scenarios	Possible reasons (causal factors)
S1	Necessary inputs are received, but the estimation algorithm fails to provide correct value of estimated position and heading (inadequate control algorithm), which results in wrong estimated position and heading	1) The specified control algorithm is flawed (software failure), e.g., parameters are not tuned sufficiently, leading to incorrect navigation calculation 2) Navigation error is not well handled in the algorithm, leading to unacceptable calculation accuracy of the position
S2	Position and heading of AUV is not correctly estimated since the position of buoys is not updated successfully (delayed)	1) Buoy is out of the acoustic range due to ice drift—ice drift is not correctly estimated by operator before operation, which causes the navigation buoys to be placed in the wrong position 2) Failure of acoustics of navigation buoy, leading to the failure of navigation 3) Failure of acoustic module in AUV, leading to the failure of navigation 4) GPS of navigation buoy fails to provide accurate position due to electromagnetic interference or atmospheric conditions
S3	Position and heading of AUV is not correctly estimated because the correct depth or altitude of the AUV is not received. As a result, the dead reckoning technique cannot provide correct navigation estimation	1) Failure of Conductivity, Temperature, Depth (CTD) sensor to collect depth data 2) Failure of Acoustic Doppler Current Profiler (ADCP)/Doppler Velocity Logs (DVL) to collect AUV speed 3) Failure of Inertial Measurement Unit (IMU)
S4	Position and heading of AUV is not correctly estimated because the correct speed of the AUV is not received. As a result, the dead reckoning technique cannot provide correct navigation estimation	1) Failure of ADCP/DVL to collect AUV speed. 2) Incorrect information/feedback of the propeller's rpm
S5	Position and heading of AUV is not correctly estimated because navigation system module fails to accurately measure the distance between the vehicle and navigation buoy	1) Failure of acoustic module in AUV, leading to the failure of navigation 2) High ambient noise, leading to the failure of navigation 3) Multipath from ice, leading to the failure of navigation

Table 11
Applicability of different methods to the online risk modeling of AMS.

Criteria	PHA	Procedural HAZOP	STPA
C1	Y	P	Y
C2	P	P	Y
C3	I.I.	I.I.	I.I.
C4	Y	P	Y
C5	Y	Y	Y
C6	I.I.	I.I.	Y
C7	P	Y	Y
C8	P	P	Y
C9	N	N	N
C10	L	L	H
C11	N	N	N
C12	P	P	P
C13	Y	Y	Y
C14	N	N	N
C15	N	N	N

Abbreviations: Y-Yes, N–No, P-Partial, I.I.-Insufficient information, L-Low, H-High.

traced back to the failure of the Conductivity, Temperature, Depth (CTD) sensor, ADCP/Doppler Velocity Logs (DVL), and Inertial Measurement Unit (IMU), as shown in Table 10. However, since physical components are not explicitly described and shown during the analysis, such as in PHA, identifying physical failures might not be as easy as in PHA, though similar results in terms of the reliability and maintenance aspects are obtained in the current study.

Compared to the other two methods, Procedural HAZOP does not provide a satisfactory result in terms of the reliability and maintenance aspects of the system, since the current Procedural HAZOP mainly focuses on procedures or operational sequences. Though some of the reliability-related issues can be identified, for example, a sensors issue is identified as a cause leading to the status of the AUV being incorrectly reported during the mission, and hardware failures, such as transmitter and drive board, are identified as causes leading to the failure of testing communication before operation, as shown in Table 8, Procedural HAZOP fails to provide a more detailed analysis.

7.2.2. Inclusion of the performance of software and control algorithm

For the PHA, a simple checklist-based method was employed. The level of detail of the analysis results sometimes depends on the checklist. Without a detailed checklist on software failure and interaction among components, PHA may not be sufficiently detailed to analyze the software and control algorithms' performance. Taking the hazardous event *Failure of self-designed software in AUV (No. 31)* in the PHA results, for example, the possible causes including a bad configuration and possible bugs caused by untested features, software updates, compliance issues, and insufficient functional design are identified. Apparently, the identified causes provide a general idea of how the software can fail, but without a detailed checklist on software failure, it is difficult to further refine both the hazardous event and its causes based on analysts' experience alone.

In the case of Procedural HAZOP, the current analysis mainly focuses on the operational procedure, in which the deviation is mainly related to operators' behavior. Though software failures could be identified as the causes of possible deviations, for example, the possible configuration issues leading the failure of testing communication before operation or failed detection of critical error of AUV, it is not easy to perform a more detailed analysis on how these failures occur since there are no specific guidewords on software failures to facilitate it. Other studies, however, claim that the combination of traditional HAZOP, human factor HAZOP, and software HAZOP might help to identify more software-related hazards, though further work is required (Sultana et al., 2019).

Compared to traditional methods, STPA needs a hierarchical control structure to demonstrate the system's interaction. By breaking down the entire system and identifying the relationships in this way, it is easier to identify the software failure by analyzing how UCAs can occur. As shown in Tables 9 and 10, a hazardous event may occur if the navigation system module provides the unacceptable inaccurate estimated position and heading during the mission. Though necessary inputs are received by the routing and planning module, failure can still occur when the estimation algorithm fails to provide a correct value of the estimated position and heading. Further reasons can be identified as either flawed control algorithms, such as untuned parameters in the code, or the algorithm's inability to handle navigation error well. This analysis process is guided in the STPA method when developing a loss scenario by analyzing how the process model and feedback can lead to the potential loss.

7.2.3. Inclusion of the performance of the interaction between software and hardware

System failures can occur not just because of pure hardware failure or software failure. The performance of the software sometimes depends on the hardware. Considering that the hardware may change with the impact from an environment, the complexity of a system with both software and hardware may increase. The failure caused by the interaction between software and hardware, such as the physical damage of hardware components caused by the electronic stress induced by software execution, also needs to be identified (Feng et al., 2014; Zhu and Pham, 2019). Failure caused by the interaction between software and hardware is not easy to identify due to few historical records and previous experience. The current analysis results do not identify any related hazards, but this may be because of the limited experience and knowledge of the analysis group. A detailed analysis of this aspect should be conducted in the future.

7.2.4. Inclusion of the performance of the interaction between AMS and external supporting system

In the current study, only the navigation buoys system is considered as an external supporting system for the AUV operation. In the current AUV under-ice operation, the navigation buoys system is essential to support the AUV navigation and operators' control and monitoring. As shown in the PHA results, hazards related to the AUV and navigation buoys are analyzed separately. Though the interaction between the AMS and the external supporting system is not explicitly described in PHA, the method considers this issue in another way. For example, navigation buoys drifting out of the operation area due to strong wind or current (No. 58 and No. 61) are identified when analyzing the risk caused by environmental hazards.

A similar risk is identified in the interaction between the AUV and navigation buoys system in STPA. A buoy out of the acoustic range due to ice drift causes navigation buoys not to be appropriately placed. This thus causes the position of the buoys not to be updated successfully, leading to the occurrence of *Navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission* (UCA5-P-1). Though similar results might be obtained in PHA and STPA, visualizing the interaction between the AMS and the external supporting system in STPA's control structure and analyzing the interaction explicitly provides operators with a clearer view. The visualization of the interaction between the AMS and the external supporting system is very valuable in the analysis of AMS operation, in which these interactions might bring more issues compared to the operation of traditional marine systems, and therefore require special attention from operators.

7.2.5. Inclusion of the performance of the communication between the AMS and environment

Environmental hazards are explicitly described as one of the main hazards in PHA, as shown in Table 6. Possible environmental hazards identified in PHA include the strong wind, water current, existence of ice, temperature and salinity of water, wave, etc. At the same time, the system boundary used for STPA analysis usually includes the parts of the system over which the system designers have some control, according to the STPA Handbook (Leveson and Thomas, 2018). Therefore, the environment is not directly regarded as part of the system due to its uncontrollability in the current analysis, and environmental hazards are indirectly identified as the causal factors of UCAs in STPA. Since the under-ice operation is planned in the Arctic, analyzing and listing environmental hazards explicitly might be better for operators concerned with the environmental impact on the systems. Obviously, this should be taken into consideration for the operation in harsh environments when performing risk analysis.

Procedural HAZOP does not perform well when directly analyzing the interaction between the AMS and the environment. However, compared to the other two methods, it provides a much more complete and detailed analysis of how the environment affects the functioning of operators, which might then affect the operation and performance of the AUV. For example, during the deployment stage, the weather condition may cause operators to have insufficient time to deploy the AUV or navigation buoys, which in turn may lead to reduced mission time or reduced navigation coverage of buoys. This hazardous event, captured by Procedural HAZOP, is ignored in PHA and STPA. It can be shown that, compared to PHA or STPA, Procedural HAZOP can provide some new ideas in terms of the environmental hazards.

7.2.6. *Inclusion of the hazards and possible changes in risk models caused by adaptive autonomy/mode or the change of subsystems involved*

The current AUV under-ice operation task does not involve the adaptive autonomy/operation mode, so argument using analysis results is impossible in the present study. However, previous studies may provide some ideas. Yang et al. propose a systems-theoretic approach based on STPA to deal with the marine system with dynamic autonomy, in which the transition between two modes is emphasized. The possible UCAs due to the transition are analyzed by adapting four ways in which a control action can be unsafe in STPA (Yang et al., 2020b), demonstrating the applicability of STPA in terms of adaptive autonomy. PHA and HAZOP might be able to be adapted to consider this; however, little research has been done on this so far.

7.2.7. *Inclusion of the human–machine interaction*

As shown in the results, eight hazardous events related to human operators are identified in the PHA analysis, mainly focusing on task planning, maintenance and testing of both hardware and software, remote monitoring and control during the operation, and system design. However, similarly to the software failure, it is challenging to further refine the analysis without a specific hazard list related to human–machine interaction. STPA describes operators as part of the control structure system, as shown in Fig. 3, and the interaction is represented as control actions and feedback. Compared to the results obtained from PHA, STPA performs much better in identifying and analyzing the interaction between operators and the navigation buoys system. During operation, the navigation buoys system is used to transfer some control actions from the operators to the AUV and feedback from the AUV to the operators. The PHA results ignore this intermediate system when analyzing the interactions between the operators and the AUV. In contrast, the interaction between the operators and the AUV and the interaction between the operators and the navigation buoys system are separately represented in STPA, providing a more detailed analysis.

With the help of a detailed operational procedure, different operation phases from pre-deployment to post-deployment are considered in Procedural HAZOP. By analyzing the way the operations deviate from the designed procedure, it is possible to identify how the human operators' behavior affects the AUV operation, and how the issues in the AUV affect the decision-making of the human operators and thus the subsequent operations. For example, when operators are unfamiliar with the operational procedure, it may take a longer time to test or deploy AUV and buoys than expected, which can cause a delayed mission and a possible unacceptable collected data due to limited operation time. On the contrary, if AUV gets stuck under ice or one of buoys is frozen in the ice, human operators need longer time to recover the AUV or buoys due to increased difficulties. The subsequent mission may then be affected due to limited operation time or the damaged AUV or buoys from the difficult recovery.

It is found that focusing on the operational procedure, the Procedural HAZOP results provide a more detailed view of human behavior, which neither PHA nor STPA covers well. In addition, as claimed by Utne and Schjølberg (2014), certain phases of AUV operations in the Arctic may involve a higher level of risk to humans, such as during the launch and recovery of the vehicles. Procedural HAZOP separates the operation into several phases, providing a clearer view of the hazardous events during each phase. This may make it the preferred method for those operators who are concerned with certain dangerous operational phases. However, challenges may arise if the designed procedure is not valid for the operation. The hazardous event caused by this might be difficult to identify during the analysis.

7.2.8. *Inclusion of the security issue*

In the current study, none of the three analysis methods considers security. This is not due to these methods' inability, but because of the research scope and type of operation in the current case study. A Security Vulnerability Analysis (SVA) is quite similar to a PHA and HAZOP, in terms of procedure and documentation. An SVA evaluates risk from deliberate acts resulting in accidents or incidents. Thus, an existing PHA or HAZOP can be efficiently and effectively expanded to add SVA to include the possible security issue (Nolan, 2014). An example can be found in Thieme et al. (2019), where security and cybersecurity are included in a PHA of an auto-ferry operation. An extension of STPA, called STPA-SEC, was proposed to solve the security issue (Leveson, 2004). Several studies have been conducted using STPA-SEC for security analysis (Sayers et al., 2020; Schmittner et al., 2016; Sidhu, 2018), and the results have proved its validity in solving security issues. A method's ability to solve security issues is also related to its ability to assess the software and control algorithms' performance (see Sections 7.2 and 7.3).

7.2.9. *Inclusion of various sources of data to estimate the risk level*

All three methods gather different knowledge in the analysis, including historical data, expert experience, and specific design information of the current AUV and navigation buoys system. Still, none of them provide real-time quantitative estimations of the risk level, which does not make use of sensor data. However, they can help identify which data should be collected and utilized to construct an online risk model. For example, monitoring the distance between navigation buoys during the operation might be necessary to improve the safe operation of the AUV. This conclusion can be derived from the hazardous events No. 58 and No. 61 from the PHA analysis. Hazardous events identified in Procedural HAZOP, such as *Operators do not/unable to test acoustic communication between buoys and AUV*, also show the necessity to monitor the distance between navigation buoys during the operation. A similar conclusion can be drawn from several of the UCAs identified in STPA, including UCA5-P-1 (Navigation system module provides (unacceptable) inaccurate estimated position and heading during the mission), UCA17-N-1 (Navigation system module of AUV does not measure the range between the vehicle and the buoy when the vehicle is operating under water), UCA18-N-1 (Navigation buoys system does not provide the position of buoys to AUV when the vehicle is operating under water), UCA19-N-1 (AUV operator does not stop the mission when the AUV mission is found to have failed) and so on.

7.2.10. *Level of knowledge needed for analysis*

The PHA should be carried out by those who have a background in safety engineering, and it requires experience and understanding of the system. A HAZOP study is carried out as several brainstorming sessions by a group of experts. Compared to PHA and HAZOP, the development of STPA requires more experience with the method. For example, the STPA analysis results rely heavily on the quality of the control structure in step 2. The development of the control structure depends on the analyst's knowledge of the system and the ability to conceptualize the system. In general, STPA needs sufficient knowledge to build the structure and identify UCAs. Based on the experience in the current study, both PHA and HAZOP were less time-consuming than STPA. When knowledge and time are limited, PHA or HAZOP are probably good choices with acceptable analysis results.

7.2.11. Be able to update risk level with new information/data

All these three methods are qualitative or semi-quantitative risk assessment methods. PHA and HAZOP may use a risk matrix to generate a basic ranking of risk values, providing operators with information about which part of the system or indicator is more important than others, as well as the information that needs more attention for monitoring. This can help risk analysts determine what kind of new data should be collected and then to further develop a dynamic model based on these indicators, thereby updating the risk level with new information.

One of the limitations of STPA is that it cannot provide a quantitative risk measurement. Though it might provide more nearly complete analysis in terms of software failure and the interaction among components as discussed in previous sections, operators might be overwhelmed by a long list of loss scenarios without knowing their severity. Therefore, it is challenging to know which information should be prioritized for monitoring and then updating.

7.2.12. Be able to deal with emerging risks

Though various definitions are given in different studies, the concept of emerging risk is usually associated with new (types of) events and related to known unknowns (Flage and Aven, 2015). In the operation of an AMS, emerging risks might occur due to several factors, including (Wróbel et al., 2018a):

1) The level of detail of the analysis is relatively low when the operational experience is limited, such as the operation of MASS (Chaal et al., 2020; Wróbel et al., 2018a, b) and under-ice operation of AUVs, causing only general statements to be made.
2) The complexity of the system and the nature of the interaction between its components lead to multidirectional failure propagation that analysts cannot identify.

PHA and STPA can be used to evaluate hazards early in a project being undertaken at the conceptual stage and can be refined later through additional and more thorough studies. For example, as more details of the AUV or navigation buoys system in this study are accessible, a more detailed list of components can be provided when analyzing technical hazards in PHA, and a more detailed control structure in STPA can be used to obtain a more detailed description of the control actions and feedback. In terms of the complex interaction in the system, the ability to handle this issue has been discussed in previous subsections. Hence, it can be concluded that STPA can provide a better analysis than PHA and HAZOP.

All three methods, however, are unable to guarantee that all potential hazardous events and scenarios can be addressed. A general challenge with hazard identification is that there is no or little feedback. Analysts might miss an unidentified hazard until it occurs, and the consequences may turn out differently. The monitored data is expected to provide some pre-warning or feedback to the risk analyst, helping to deal with emerging hazards.

7.2.13. Be able to efficiently identify RIFs that need to be monitored online or in real time during operation

Results from the three methods can be used to identify RIFs that can be used in the development of the online risk model. Though monitored RIFs can provide a measurement of an online risk value, it is almost impossible to monitor all identified RIFs due to challenges with quantification and costs. Determining which RIFs should be prioritized is therefore essential.

As discussed in Section 7.11, compared to PHA and HAZOP, STPA might provide an overwhelming list of input to deriving RIFs that operators cannot easily handle due to the difficulties in ranking the importance of loss scenarios. A specific risk model focusing on the RIFs that are derived from several UCAs of interest, such as the model developed by Utne et al. (2020), might be solvable; however, providing the overall risk spectrum of an AMS using the RIFs derived from the full list of UCAs in STPA can be challenging.

7.2.14. Be able to effectively model the correlation between identified RIFs

None of the three methods quantitatively describes the correlation among RIFs like FTA or BBN. However, the analysis logic behind these methods might be able to help quantitatively identify the correlation among them in constructing an online risk model in the next stage.

In STPA, the results are provided in a top-down manner. The UCAs identified are used to develop loss scenarios, and a more detailed analysis can be performed to refine the possible loss scenarios. Rokseth et al. (2018) represent this refinement process of loss scenarios as a tree structure. Through this tree structure, the relationship among RIFs can be preliminarily determined. This top-down process of STPA results can be transformed and represented using FTA or BBN, as demonstrated in previous studies by Bolbot et al. (2020) and Utne et al. (2020). Given the preliminary relationship determined in STPA, statistics or expert judgment can be used to further determine the detailed correlation among identified RIFs in the online risk model.

Developing a quantitative risk model based on PHA and HAZOP requires analysts to fully understand the hazards and their potential effects from the analysis. It is not easy to determine the correlation among identified RIFs without a structured way to show the results, which may be easier with the control hierarchy in STPA. Even with a list of hazardous events including causes and consequences as in PHA, analysts still need to extract the necessary information from the results and then determine the relationship among the identified RIFs.

7.2.15. Be able to deal with the uncertainty

Online risk models rely on real-time monitoring to estimate the risk level of the system. Therefore, the sensor and data fusion algorithms' uncertainty can significantly affect the accuracy of online risk models. Although the methods identify hazards associated with sensors such as the CTD sensor and navigation sensors, the existing sensors in the AMS might be insufficient to provide the whole online risk picture. Other kinds of data might be required to further construct the online risk model, and other new sensors might also be needed. The risk caused by these sensors and data fusion algorithms should be accurately measured and reflected in the future online risk model.

7.3. Verification and validation

The verification and validation of hazard identification results are always a challenging issue since the process of hazard identification and risk analysis often needs multiple iterations to improve the accuracy and completeness of the results. Also, these analyses often address incidents for which there is limited experience with. However, considering the following points, the analysis results and the derived conclusion in this study can be considered acceptable and credible to a certain extent.

Firstly, three methods used in the current study have been widely used in hazard identification and risk assessment. Many previous studies have tested their effectiveness and validity (Rokseth et al., 2017; Sultana et al., 2019). Also, PHA and HAZOP have been widely used and proved in industries such as oil and gas industry and marine and offshore industry (Rausand, 2013). The well-structured steps in these methods make it easy for them to provide reasonable results. In the current study, PHA and Procedural HAZOP are methods that based on brainstorming sessions, and STPA is also performed based on researcher's experience. The quality of the analysis mainly depends on the knowledge from different expert and historical experience. Researchers in the analysis group have a good knowledge in AUV and have many years of experience in operating AUV under different environmental conditions, which can help to provide an acceptable and credible results.

In addition, the hazard identification results from three methods demonstrate good agreement with the analysis results in other studies and historical operation and fault log reported in previous AUV operations (Brito et al., 2010; Ferguson, 2008; Kaminski et al., 2010). For example, according to the fault log reported in the study by Brito et al. (2010), the aborted mission due to bad crimp joint has been detected in a previous AUV operation, which is also identified in the current PHA results; the mission failure caused by uncertainty in indicated motor rpm is also identified in the current STPA results. Some identified hazards that specific to the AUV operation in harsh environment are also found in previous AUV operations in the Arctic. For example, the failure of equipment during deployment, such as CTD sensor, caused by the large temperature gradients in the air and underwater (Ferguson, 2008; Kaminski et al., 2010), which has been identified in the PHA and Procedural HAZOP results.

However, due to the limitations of researchers' knowledge and the shortcoming of applied methods in certain aspects, the current study cannot guarantee all hazardous events related to AUV under-ice operation are identified. Regular updates and improvements should be made to improve the accuracy and completeness of the results in the future.

In terms of the conclusion derived in Section 7.2, all the points and opinions are generalized from the analysis process of three methods. Detailed examples in the analysis results are provided in this section to prove and support the point of view from researchers. However, the changes in the input to the analysis do affect the derived conclusion. The main essential influencing factors include the available information or historical data of the studied system and analysts' knowledge of the studied system. In order to reduce the influence of the change in these factors, the same analysts were involved both in PHA and Procedural HAZOP workshops; STPA analysis was initially performed by the first author and then reviewed by the same analyst as the PHA and Procedural HAZOP workshops. In terms of the possible influences of the selected case study on the conclusions drawn, although the AUV under-ice operation is used as a case study, many attributes and characteristics are shared by other AMS. In addition, since the criteria developed in the current study are more or less generic, it is expected that the analysis results and conclusion could be adapted to other AMS as well.

8. Conclusions and future work

The current study identifies criteria that reflect the aspects that should be considered when developing an online risk model for AMS, and these criteria may also be used as a checklist to verify and improve the existing analysis results. The current work investigates how PHA, STPA, and Procedural HAZOP contribute to fulfilling the different requirements for online risk modeling of AMS, and how the results may be used as the basis for model development.

The case study in the article addresses an AUV under-ice operation in the Arctic. Considering the challenges in underwater navigation, the online risk model should mainly focus on the interaction between the AUV and its navigation buoys system and also the behavior of human operators during the operation.

Considering that most AMS have similar requirements and demands as AUVs with respect to the online risk model, the analysis results and conclusion from the current study can also be adapted to other AMS. Generally, the analysis results show that, compared to the other two methods, STPA is considered a good basis for developing an online risk model in terms of the number of criteria fulfilled, and especially its ability to handle the interaction among systems and software failure, although some disadvantages prevent it from becoming an ideal one.

A challenge with using STPA is the difficulties in determining the RIFs that should be included and monitored in an online risk model based on the exhaustive list of unranked loss scenarios. Considering its relatively good coverage of identified evaluation criteria, however, and that some of the RIFs may be difficult to measure in operation even though they are important in the case of AMS, such as the performance of software and control algorithm, it is worth considering STPA as a basis for the further development of online risk models. In addition, considering the changing role of human operators in the operation of AMS, it is necessary to identify specific RIFs related to human operators in the operation of AMS. Hence, some of the disadvantages with STPA may be mitigated by using PHA and Procedural HAZOP as complementary tools. In future works, the results obtained from these three methods will be used to develop an online risk model for the AUV operation. The criteria identified for the online risk model of AMS in the present study will also be used as a checklist to verify and improve the quality when developing the online risk model.

CRediT authorship contribution statement

Ruochen Yang: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Ingrid Bouwer Utne:** Conceptualization, Methodology, Formal analysis, Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors appreciate the funding support of the Research Council of Norway through the Nansen Legacy project with project number RCN#276730. The authors are also very grateful for the constructive feedback and comments from the anonymous reviewers to an earlier version of this article.

References

Baksh, A.-A., Abbassi, R., Garaniya, V., Khan, F., 2018. Marine transportation risk assessment using Bayesian Network: application to Arctic waters. Ocean Eng. 159, 422–436.

Barua, S., Gao, X., Pasman, H., Mannan, M.S., 2016. Bayesian network based dynamic operational risk assessment. J. Loss Prev. Process. Ind. 41, 399–410.

Blanchard, B.S., 2004. System Engineering Management. John Wiley & Sons.

Bolbot, V., Theotokatos, G., Boulougouris, E., Psarros, G., Hamann, R., 2020. A novel method for safety analysis of cyber-physical systems—application to a ship exhaust gas scrubber system. Saf. Now. 6 (2), 26.

Brito, M., Griffiths, G., 2016. A Bayesian approach for predicting risk of autonomous underwater vehicle loss during their missions. Reliab. Eng. Syst. Saf. 146, 55–67.

Brito, M.P., Griffiths, G., 2011. A Markov chain state transition approach to establishing critical phases for AUV reliability. IEEE J. Ocean. Eng. 36 (1), 139–149.

Brito, M.P., Griffiths, G., Challenor, P., 2010. Risk analysis for autonomous underwater vehicle operations in extreme environments. Risk Anal.: Int. J. 30 (12), 1771–1788.

Broadleaf, 2018. Technical Note: Process and Guidewords for Procedural HAZOPs. In: https://broadleaf.com.au/resource-material/process-and-guidewords-for-procedural-hazops/.

Chaal, M., Banda, O.A.V., Glomsrud, J.A., Basnet, S., Hirdaris, S., Kujala, P., 2020. A framework to model the STPA hierarchical control structure of an autonomous ship. Saf. Sci. 132, 104939.

Chen, X., Bose, N., Brito, M., Khan, F., Thanyamanta, B., Zou, T., 2021. A review of risk analysis research for the operations of autonomous underwater vehicles. Reliab. Eng. Syst. Saf. 216, 108011.

Feng, E., Zheng, J., Liu, C., 2014. An integrated reliability model of hardware-software system. In: 2014 10th International Conference on Reliability, Maintainability and Safety (ICRMS). IEEE, pp. 577–580.

Ferguson, J., 2008. Adapting AUVs for Use in Under-ice Scientific Missions, OCEANS 2008. IEEE, pp. 1–5.

Flage, R., Aven, T., 2015. Emerging risk–Conceptual definition and a relation to black swan type of events. Reliab. Eng. Syst. Saf. 144, 61–67.

Hegde, J., Henriksen, E.H., Utne, I.B., Schjølberg, I., 2018a. Development of dynamic safety envelopes for autonomous remotely operated underwater vehicles. Safety and Reliability–Safe Societies in a Changing World. Proc ESREL, 2018, June 17-21, 2018, Trondheim, Norway.

Hegde, J., Henriksen, E.H., Utne, I.B., Schjølberg, I., 2019. Development of safety envelopes and subsea traffic rules for autonomous remotely operated vehicles. J. Loss Prev. Process. Ind. 60, 145–158.

Hegde, J., Utne, I.B., Schjølberg, I., Thorkildsen, B., 2018b. A Bayesian approach to risk modeling of autonomous subsea intervention operations. Reliab. Eng. Syst. Saf. 175, 142–159.

IEC, 2016. Hazard and Operability Studies (HAZOP Studies) - Application Guide.

Information-technology Promotion Agency, 2021. STAMP Workbench's HomePage. https://www.ipa.go.jp/english/sec/reports/20180330.html. (Accessed 30 November 2021).

Kaminski, C., Crees, T., Ferguson, J., Forrest, A., Williams, J., Hopkin, D., Heard, G., 2010. 12 Days under Ice–An Historic AUV Deployment in the Canadian High Arctic, 2010 IEEE/OES Autonomous Underwater Vehicles. IEEE, pp. 1–11.

Khakzad, N., Khan, F., Amyotte, P., 2012. Dynamic risk analysis using bow-tie approach. Reliab. Eng. Syst. Saf. 104, 36–44.

Khakzad, N., Khan, F., Amyotte, P., 2013. Dynamic safety analysis of process systems by mapping bow-tie into Bayesian network. Process Saf. Environ. Protect. 91 (1–2), 46–53.

Khan, F., Hashemi, S.J., Paltrinieri, N., Amyotte, P., Cozzani, V., Reniers, G., 2016. Dynamic risk management: a contemporary approach to process safety management. Curr. Opin. Chem. Eng. 14, 9–17.

Kim, H., Lee, S.-H., Park, J.-S., Kim, H., Chang, Y.-S., Heo, G., 2015. Reliability data update using condition monitoring and prognostics in probabilistic safety assessment. Nucl. Eng. Technol. 47 (2), 204–211.

Laurinen, M., 2016. Remote and Autonomous Ships: the Next Steps. Advanced Autonomous Waterborne Application Partnership. Buckingham Gate, London.

Lazakis, I., Dikis, K., Michala, A.L., Theotokatos, G., 2016. Advanced ship systems condition monitoring for enhanced inspection, maintenance and decision making in ship operations. Trans. Res. Procedia 14, 1679–1688.

Leveson, N., 2004. A new accident model for engineering safer systems. Saf. Sci. 42 (4), 237–270.

Leveson, N., 2011. Engineering a Safer World: Systems Thinking Applied to Safety. MIT press.

Leveson, N.G., Thomas, J.P., 2018. STPA Handbook, 3, 1-188.

Liu, Z., Ma, Q., Cai, B., Liu, Y., Zheng, C., 2021. Risk Assessment on Deepwater Drilling Well Control Based on Dynamic Bayesian Network. Process Safety and Environmental Protection.

Loh, T.Y., Brito, M.P., Bose, N., Xu, J., Tenekedjiev, K., 2019. A Fuzzy-Based Risk Assessment Framework for Autonomous Underwater Vehicle Under-Ice Missions. Risk Analysis.

Nolan, D.P., 2014. Safety and Security Review for the Process Industries: Application of HAZOP, PHA, What-IF and SVA Reviews. Elsevier.

Norgren, P., Mo-Bjørkelund, T., Gade, K., Hegrenæs, Ø., Ludvigsen, M., 2020. Intelligent buoys for aiding AUV navigation under the ice. In: 2020 IEEE/OES Autonomous Underwater Vehicles Symposium (AUV)(50043). IEEE, pp. 1–7.

Paltrinieri, N., Scarponi, G., Khan, F., Hauge, S., 2014. Addressing dynamic risk in the petroleum industry by means of innovative analysis solutions. Chem. Eng. Trans. 36, 451–456.

Rausand, M., 2013. Risk Assessment: Theory, Methods, and Applications. John Wiley & Sons.

Rokseth, B., Utne, I.B., Vinnem, J.E., 2017. A systems approach to risk analysis of maritime operations. Proc. Inst. Mech. Eng. O J. Risk Reliab. 231 (1), 53–68.

Rokseth, B., Utne, I.B., Vinnem, J.E., 2018. Deriving verification objectives and scenarios for maritime systems using the systems-theoretic process analysis. Reliab. Eng. Syst. Saf. 169, 18–31.

Rosness, R., 1998. Risk influence analysis a methodology for identification and assessment of risk reduction strategies. Reliab. Eng. Syst. Saf. 60 (2), 153–164.

Sayers, J.M., Feighery, B.E., Span, M.T., 2020. A STPA-sec case study: eliciting early security requirements for a small unmanned aerial system. In: 2020 IEEE Systems Security Symposium (SSS). IEEE, pp. 1–8.

Schmittner, C., Ma, Z., Puschner, P., 2016. Limitation and improvement of STPA-Sec for safety and security co-analysis. In: International Conference on Computer Safety, Reliability, and Security. Springer, pp. 195–209.

Sidhu, A.S., 2018. Application of STPA-Sec for Analyzing Cybersecurity of Autonomous Mining Systems. Massachusetts Institute of Technology.

Sultana, S., Okoh, P., Haugen, S., Vinnem, J.E., 2019. Hazard analysis: application of STPA to ship-to-ship transfer of LNG. J. Loss Prev. Process. Ind. 60, 241–252.

The Nansen Legacy. https://arvenetternansen.com/(accessed 31 May 2021).

Thieme, C.A., Guo, C., Utne, I.B., Haugen, S., 2019. Preliminary hazard analysis of a small harbor passenger ferry–results, challenges and further work. In: Journal of Physics: Conference Series. IOP Publishing, 012024.

Thieme, C.A., Utne, I.B., 2017. Safety performance monitoring of autonomous marine systems. Reliab. Eng. Syst. Saf. 159, 264–275.

Thieme, C.A., Utne, I.B., Haugen, S., 2018. Assessing ship risk model applicability to marine autonomous surface ships. Ocean Eng. 165, 140–154.

Utne, I.B., Rokseth, B., Sørensen, A.J., Vinnem, J.E., 2020. Towards supervisory risk control of autonomous ships. Reliab. Eng. Syst. Saf. 196, 106757.

Utne, I.B., Schjølberg, I., 2014. A Systematic Approach to Risk Assessment: Focusing on Autonomous Underwater Vehicles and Operations in Arctic Areas. ASME 2014 33rd International Conference on Ocean, Offshore and Arctic Engineering. American Society of Mechanical Engineers Digital Collection.

Vagia, M., Transeth, A.A., Fjerdingen, S.A., 2016. A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? Appl. Ergon. 53, 190–202.

Vinnem, J.-E., Røed, W., 2019. Offshore Risk Assessment, 2. Springer.

Vinnem, J.E., Utne, I.B., Schjølberg, I., 2015. On the need for online decision support in FPSO–shuttle tanker collision risk reduction. Ocean Eng. 101, 109–117.

Wang, Y.F., Qin, T., Li, B., Sun, X.F., Li, Y.L., 2017. Fire probability prediction of offshore platform based on Dynamic Bayesian Network. Ocean Eng. 145, 112–123.

Wróbel, K., Montewka, J., Kujala, P., 2017. Towards the assessment of potential impact of unmanned vessels on maritime transportation safety. Reliab. Eng. Syst. Saf. 165, 155–169.

Wróbel, K., Montewka, J., Kujala, P., 2018a. System-theoretic approach to safety of remotely-controlled merchant vessel. Ocean Eng. 152, 334–345.

Wróbel, K., Montewka, J., Kujala, P., 2018b. Towards the development of a system-theoretic model for safety assessment of autonomous merchant vessels. Reliab. Eng. Syst. Saf. 178, 209–224.

Yang, R., Utne, I., Liu, Y., Paltrinieri, N., 2020a. Dynamic risk analysis of operation of the autonomous underwater vehicle (AUV). In: The 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference, Italy.

Yang, X., Utne, I.B., Sandøy, S.S., Ramos, M.A., Rokseth, B., 2020b. A systems-theoretic approach to hazard identification of marine systems with dynamic autonomy. Ocean Eng. 217, 107930.

Zadakbar, O., Khan, F., Imtiaz, S., 2015. Dynamic risk assessment of a nonlinear non-Gaussian system using a particle filter and detailed consequence analysis. Can. J. Chem. Eng. 93 (7), 1201–1211.

Zeng, Z., Zio, E., 2018. Dynamic risk assessment based on statistical failure data and condition-monitoring degradation data. IEEE Trans. Reliab. 67 (2), 609–622.

Zhu, M., Pham, H., 2019. A novel system reliability modeling of hardware, software, and interactions of hardware and software. Mathematics 7 (11), 1049.

Zio, E., 2018. The future of risk assessment. Reliab. Eng. Syst. Saf. 177, 176–190.