Ocean Engineering 217 (2020) 107930

Contents lists available at ScienceDirect

Ocean Engineering

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

A systems-theoretic approach to hazard identification of marine systems with dynamic autonomy

Xue Yang *, Ingrid B. Utne, Stian S. Sandøy, Marilia A. Ramos, Børge Rokseth

Norwegian University of Science and Technology, Department of Marine Technology, 7491 Trondheim, Norway

ARTICLE INFO

Keywords:
Hazard identification
Dynamic autonomy
Level of autonomy
STPA
Transitions

ABSTRACT

Autonomous marine systems may switch between various operational modes with different levels of autonomy (LoA), due to a rapidly changing environment and the complex nature of tasks. The dynamic autonomy brings an additional layer of complexity to ensuring safe marine operations, but this functionality is not sufficiently considered in current risk analysis methods. Hence, this paper proposes an approach to hazard identification based on the system theoretic process analysis (STPA) that includes unsafe transitions between different LoA in systems. A case study of a remotely operated vehicle (ROV) with four operational modes with different LoAs is used to illustrate the approach. The results show that the proposed approach contributes to: 1) communicating a shift of responsibilities among human operator and system controller in different operational modes by specifying how the allocation of the responsibility between human operators and the controller changes, and what updated process model of the operator and the controller are to ensure a successful transition; 2) refining safety constraints to be more concrete to improve system design, and operational procedures and 3) identifying triggering events for marine system modes' transitions to handle environmental interaction systematically and sufficiently.

1. Introduction

Technological developments in software and hardware have led to a rapid increase in autonomous functionality in several systems and applications. Examples include transportations systems, such as autonomous cars, ships and trains, and systems for research in harsh, remote environments for reducing human exposure (Fan et al., 2020; Ramos et al., 2019a). A desired outcome of autonomy is the development of systems that operate in a more cost-effective and safe manner.

Autonomy means that the system has the "ability of integrated sensing, perceiving, analysing, communicating, planning, decision-making and acting to achieve the goals assigned by human operators through designed human-machine interface" (Utne et al., 2017a). Autonomous systems may have different levels of autonomy (LoA), and there are different classifications of LoAs. In this paper, we adopt the definition from Ludvigsen and Sørensen (2016) and Utne et al. (2017b), which classifies autonomous operations into four levels: (i) automatic operation (remote operation), (ii) management by consent, (iii) semi-autonomous or management by exception, and (iv) highly autonomous.

Autonomous marine systems may switch between various operational modes with different LoA due to rapidly changing environment or complex nature of tasks. This means that the operation, for example, may start in a lower level of autonomy (e.g. remote control mode) by a human operator, evolves into a higher level of autonomy (e.g. semi-autonomous mode) with the operator then acting as a supervisor, and later return to a lower level again in a subsequent moment, due to operational requirements and changing external conditions. For instance, dynamic LoA is expected for underwater vehicles (Ludvigsen and Sørensen, 2016; Schjølberg and Utne, 2015; Sørensen and Ludvigsen, 2015) and autonomous ships (Ramos et al., 2019b; Wu et al., 2020). The changes in LoA resulted from switching between operational modes may take place in a very short period of time (e.g. minutes) and under severe sea conditions.

The development process of autonomous marine systems must consider the risks involved in its operation to ensure safety for humans, minimal negative impact on the environment, and asset integrity (Utne et al., 2017b). For autonomous ships expected to come into operation within a few years, the efforts have been on identifying high-level hazards originating from both design and operation (Burmeister et al., 2014; Heikkilä et al., 2017; Rødseth and Burmeister, 2015; Wróbel et al.,

* Corresponding author.
E-mail address: xue.yang@ntnu.no (X. Yang).

https://doi.org/10.1016/j.oceaneng.2020.107930
Received 24 September 2019; Received in revised form 7 August 2020; Accepted 8 August 2020
Available online 11 September 2020

Abbreviations

APS	Acoustic positioning systems
DP	Dynamic positioning
DVL	Doppler velocity log
GUI	Graphical user interface
HAZID	HAZard IDentification
LoA	Level of autonomy
MCM	Manual control mode
NTM	Net pen tracking mode
ROV	Remotely operated vehicle
RPM	Revolutions per minute
RSKM	Relative station keeping mode
SC	Safety constraints
SKM	Station keeping mode
STAMP	Systems-Theoretic Accident Model and Processes
STD	State transition diagram
STPA	Systems-Theoretic Process Analysis
UCA	Unsafe control action
UTCA	Unsafe transition control actions

2016). For underwater vehicles, risk and reliabilities analyses have so far primarily focused on technical failures (Harris et al., 2016; Hinz et al., 2010; Thieme et al., 2015; Xiang et al., 2017; Xinqian et al., 2009; Xu et al., 2013), in addition to human failures to a limited extent (Ho et al., 2011; Thieme and Utne, 2017). Yet, dynamic autonomy brings an additional layer of complexity to the systems and operations, especially regarding the interactions among human operators, software, hardware. Interaction associated hazards may lead to accidents if not well recognized and controlled. Still, no risk analysis method has thus far considered how dynamic autonomy as a *functionality* can be handled in the hazard identification process.

The objective of this paper is to bridge this gap by proposing an approach to identifying hazards and safety requirements for safe operational mode transitions. Traditional hazard identification methods, such as HAZard IDentification (HAZID), checklists, and accident investigation reports, are often a basis for identifying what can go wrong, which is the first step in risk analysis (Rausand, 2011). The Systems-Theoretic Process Analysis (STPA) is a relatively new hazard analysis technique based on the causality model called the Systems-Theoretic Accident Model and Processes (STAMP) (Leveson, 2011). It defines safety as a control problem, which makes it desirable for complex systems. The STPA has been applied to identify and analyse hazards in several different domains, including autonomous ships (Valdez Banda et al., 2019; Wróbel et al., 2018b) and dynamic positioning (DP) systems on maritime vessels (Rokseth et al., 2017, 2018). Yet, to this moment system dependencies resulting from shifts in LoA has not been addressed. This paper proposes an approach that uses the STPA as a foundation and further expands it for autonomous functionality, with a particular focus on dynamic LoA resulted from mode shifting in operation. The proposed approach can contribute to the safe design of autonomous marine systems, such as underwater vehicles.

The paper is structured as follows: the approach is described in Section 2. In Section 3, the application of the approach is illustrated with a case study for the net inspection with a remotely operated vehicle (ROV) of a fish cage in aquaculture, which is followed by a discussion of the results in Section 4. Section 5 concludes the work.

2. Methodology

In STPA, there are three basic constructs: 1) *safety constraints;* 2) *hierarchical control structures;* and 3) *process models*. The safety constraints specify system conditions or behaviours that need to be satisfied to prevent hazards (Leveson and Thomas, 2018). The safety constraints should be reinforced through the behaviour of the system by an effective control system. The hierarchical control structure is "a system model that is composed of feedback control loops" (Leveson and Thomas, 2018). The process model represents internal beliefs of the controller(s) concerning the process being controlled, as well as other relevant aspects of the system or the environment. The process model thus reflects how the controller(s) perceive the system variables and their current states, the relationship between the variables and the way the process can change the states.

STPA analysis aims to determine how the safety constraints can be violated or insufficiently enforced to eliminate, mitigate and control the emergence of hazards that may develop into accidents.

For a system that has operational modes with different LoAs, smooth and safe transitions are critical for successful commission. In other words, operators must be able to take over control when needed. One merit of STPA is that it investigates dysfunctional controller interactions as a cause for flawed process execution. A regular STPA, however, does not provide explicit guidance to analysis of possible failures during operational transitions and associated shifts between different LoA. These failures are the focus of the proposed approach, as shown in Fig. 1. A color-coding is used in the figure, where there are standard procedures are marked as gray, and added procedures are marked white. A detailed description of its phases and steps is provided in the following.

2.1. Phase 1: define the purpose of the analysis

In a standard STPA, the first part of the analysis comprises:

- step 1.1: Identify losses (i.e., something of value to stakeholders);
- Step 1.2: Identify system-level hazards; and to identify system-level safety constraints.

The following *additional* steps are proposed to be integrated into phase 1 to aid in establishing control structures considering different responsibilities and process models (i.e., human operator and system controller's internal beliefs used to make decisions):

Step 1.3: Identify operational modes and corresponding LoAs

The human-machine interactions and cooperation are expressed by various LoAs, with each level specify a different degree to which an operation is in between fully manual performance and fully autonomous conditions (Vagia et al., 2016). The definition of LoA are subject to applications and should be selected to satisfy the system's own needs. In this paper, we use the following four LoAs defined by Ludvigsen and Sørensen (2016) and Utne et al. (2017b) for marine systems:

1. **LoA 1 - Automatic operation (remote control):** The human operator directs and controls all high-level mission planning. The environmental conditions and sensor data are presented to the operator through a Human-Machine-Interface (HMI);
2. **LoA 2- Management by consent:** The system automatically makes recommendations for missions or actions related to specific functions. The system can perform some tasks independently of human control when previously delegated by the human.
3. **LoA 3 - Semi-autonomous operation (management by exception)**: The system automatically executes mission-related functions. The human may override or change parameters, and cancel or redirect actions with defined timelines. The operator's attention is only brought to exceptions for certain decisions.
4. **Highly autonomous operation**: The system automatically executes missions or process-related functions in an unstructured environment with the ability to plan and replan the mission. The system is independent and intelligent.

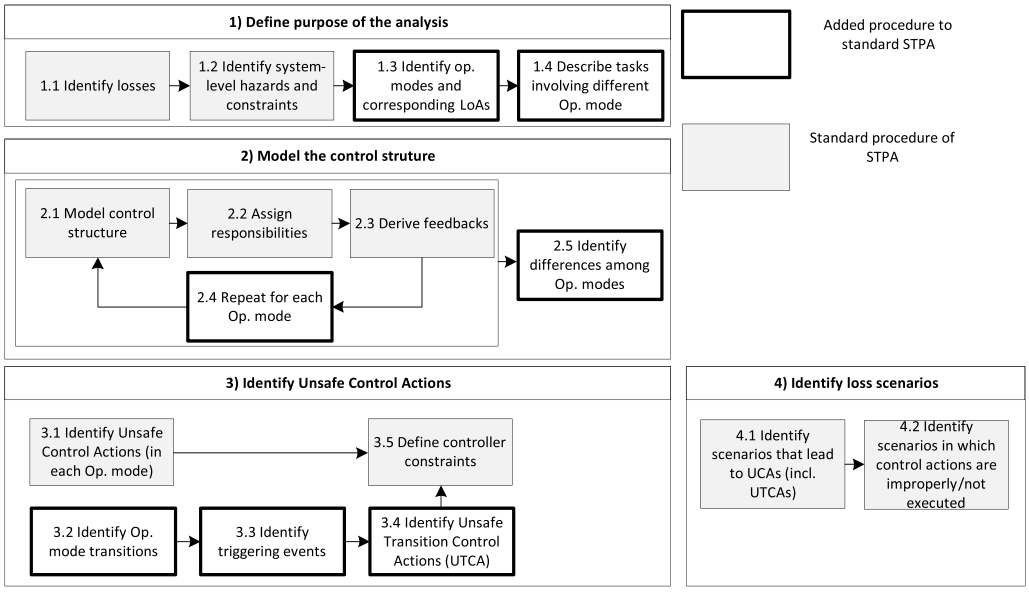

Fig. 1. Proposed approach based on STPA, including identification of Unsafe transition control actions.

In this paper, an operational mode refers to a functional configuration of a system. Inside an operational mode, "the system can perform specific operational scenarios, and so activate the corresponding functions" (Faisandier, 2013). The LoA is fixed or consistent for each operational mode. For various operational modes integrated into one marine system, the LoAs can vary across a continuum of intermediate levels between manual operation and fully autonomous.

Step 1.4: Describe tasks involving different operational modes

This step should focus on how an operation may be performed from the beginning to the end. The description does not need to be comprehensive, but to include how the controllers (i.e., both human operator and system controller) may adapt operational modes to foreseen various circumstance. The aim is to have an initial illustration of how dynamic the operation can be.

2.2. Phase 2: model the control structures

The standard STPA is functionality oriented. For each function, one control structure, rather than a physical component diagram, can be prepared at a high level and zoomed in with more details if necessary. The control structure is "a system model that is composed of feedback control loops" (Leveson and Thomas, 2018). The first three standard steps in this phase include modeling control structure, assign ing responsibilities (including process models), and deriving feedback. A repetition loop is added to analyse all the operational modes in the same manner. The control structures can help designers to understand the dynamic controller interactions associated with the shifting of LoAs to perceive, comprehend, project, and decide to complete the mission.

2.3. Phase 3: identify unsafe control actions

STPA defines an unsafe control action (UCA) as a control action that, in a particular context and worst-case environment, will lead to a hazard (Leveson and Thomas, 2018). The standard STPA procedure prescribes identifying UCAs for each operational mode (Step 3.1 in Fig. 1). In addition to the UCAs, we define unsafe *transition* control actions (UTCA). Those can be identified through the following three *additional* steps:

Step 3.2 Identify operational modes' transitions

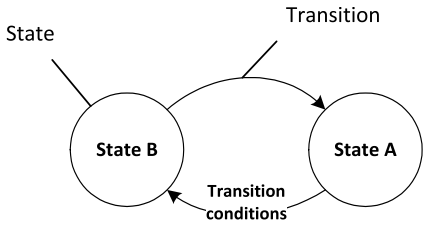

Fig. 2. Illustration of the state transition diagram.

A state transition diagram (STD) is used to analyse possible transitions between operational modes. STD was originally proposed by David Harel (1987), and is commonly applied in computer science to provide an abstract description of system behaviours. The main idea of applying STD is to demonstrate the possible transition among various operational modes with different LoA. Each transition can be investigated while analysing unsafe control action from either operator or autonomous system's controller or both. The STD is an intuitive method and perceived well represents the states of the system after shifts between finite operational modes.

Fig. 2 presents a simplified STD for possible transitions between two operational modes (states). The nodes denote the modes (states), and the arrows denote transitions. The transition happens when triggering events occur in state A, and guard conditions are satisfied (Harel, 1987).

Step 3.3 Identify triggering events

In general, the *triggering events* can be internal stimuli or external stimuli. Internal stimuli are events that occur within the autonomous system, while external stimuli are events in the surrounding operational environment. The external and internal stimuli can be described in connection with the following four types of triggering events, as adapted from (Friedenthal et al., 2014):

- A change event that happens when some condition has been satisfied;
- A timeout event that initiates after the specified amount of time elapses;

- A call event that indicates an operation has been requested by, e.g., a human operator; and
- A completion event that takes place when everything that needs to be done in the current mode is completed.

Examples of triggering events under each category are further described in detail in the case study (Section 3.3).

Step 3.4 Identify unsafe transition control actions (UTCAs)

A transition control action is an action that is either provided by a human operator or the autonomous system's controller upon transitions between modes. There are four ways a control action can be unsafe (Leveson, 2011), which also apply to unsafe transition control actions:

- not providing transition control action;
- providing a transition control action (e.g., provides a wrong transition);
- providing a transition control action too early, too late or out of order;
- the transition control action lasts too long, or is stopped too soon.

The unsafe transition control actions are defined based on the transition diagram, triggering events, responsibilities, and process models of the controllers.

2.4. Phase 4: identify loss scenarios

Once the unsafe transition control actions are identified, the next step is to identify loss scenarios and refine safety constraints if necessary. Two types of loss scenarios should be considered: (1) the scenarios that lead to a UTCA, and (2) scenarios in which the transitions are improperly executed or not executed. The first type may involve failures related to the controller, an inadequate control algorithm, an unsafe control input, and an insufficient process model (Leveson, 2011). The scenarios concerning an improper or no execution of a transition (type 2) involve a control path, which transfers transition control action to the controlled process. These scenarios might include: i) a transition control action not executed; ii) a transition control action improperly executed; and iii) the controlled process does not respond or responds incorrectly though transition control actions received (Leveson and Thomas, 2018).

3. Case study

A case study has been performed to test the feasibility of the proposed approach. The case study focuses on preventing fish escape from fish cages for aquaculture in ROV operations. Fish escape in aquaculture has severe environmental effects, and the authorities continuously put much effort to avoid escapes and mitigate the impact of escapes (Holen et al., 2019; Norwegian Ministry of Trade Industry and Fisheries, 2017).

According to the statistics for fish escape from 2010 to 2016 in sea-based aquaculture operations, most fish escapes (72%) are due to holes in the net (Føre and Thorvaldsen, 2017). The mandatory net inspection after aquaculture operations is one of the primary measures to reduce the risk of fish escape (Yang et al., 2020). The aim is to discover the holes caused by operations and to initiate necessary recovery actions as early as possible.

ROVs are prevalent in aquaculture operations, such as net inspections, but also generally to increase operation regularity, reduce exposure of the workers to the harsh environment, and increase the weather window for operations (e.g., delousing, net cleaning). ROVs have tethers that limit their manoeuvrability and increase the risk of entanglement. Nevertheless, they have a transparent and shared control process (i.e., the human controller shares control with an autonomous controller), which increases their potential for use in specific applications and environments. For instance, the shared control provides advantages when handling flexible structures, demanding environments with currents and large waves, and changing geometry in an undetermined pattern. The potential losses for such operations are primarily the loss of life due to human overboard, the loss of fish, and damage to property.

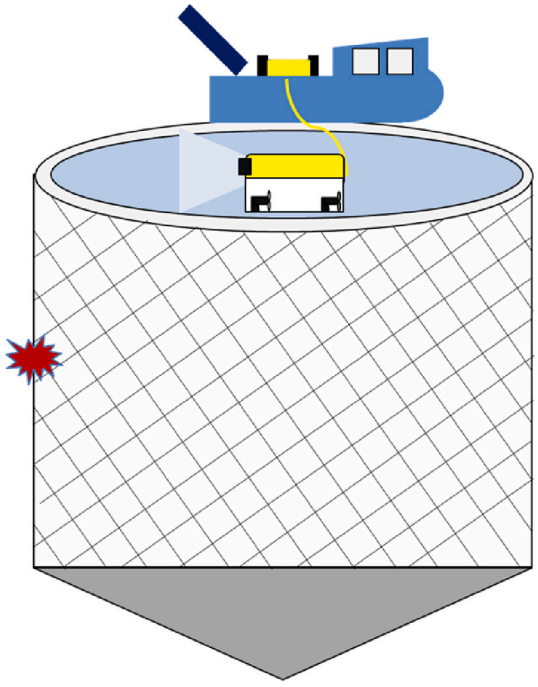

Fig. 3. Illustration of a net inspection operation in a fish cage.

The ROVs currently adopted in the industry mainly have low LoA, i.e., they are remote-controlled by an operator. ROV operation could benefit from more autonomous functionalities to increase inspection quality and effectiveness and reduce operator fatigue (Bjelland et al., 2015). Industrial (Sperre ROV Technology, 2017) and research (Schjølberg and Utne, 2015; SINTEF Ocean, 2016) efforts have been focusing on developing higher-level autonomous ROV systems (Christ and Wernli, 2014).

The proposed approach is applied to an observation class ROV provided by a vendor with extensive offshore ROV operating experience in the oil and gas industry. It engages in developing autonomous functionalities for aquaculture operations. The ROV is used for net inspection inside the fish cage (Fig. 3). The steps described in the previous section are applied to this operation in the following.

3.1. Phase 1: define the purpose of the analysis

Step 1.1: Identify losses

The main concerning loss in ROV net inspection in this case study is the loss of fish, which will result in loss of profits, loss of reputation, and biological degradation of the wild fish stocks.

Step 1.2: Identify system-level hazards and constraints

We primarily focus on the loss of fish due to: i) failures of the ROV to detect existing holes in the net; and ii) damage to the net caused by the ROV itself during operation.

The scope of the analysis is limited to two system-levels hazards:

- H-1: ROV is unable to successfully complete the inspection (incl. holes not detected);
- H-2: ROV collides with net structure or tangles with the net.

Additional system-level hazards, such as the ROV leak of hydraulic oil, the high voltage electrical hazard, frame integrity lost, and water leakage into ROVs, are not considered in this case study. Table 1 describes the safety constraints that the system conditions or behaviours need to satisfy to prevent hazards and ultimately prevent losses.

The ROV system comprises the ROV, the handling system, the surface control system, and all associated equipment. The following four ROV operational modes are considered in this case study: manual control mode (MCM), station keeping mode (SKM) and relative station keeping mode (RSKM), and net-pen tracking mode (NTM). The ROV should be capable of operating in multiple levels of autonomy to ensure operational performance.

Manual control mode (LoA 1)

In this mode, the ROV operator has direct control of each thruster via operation control (e.g., joysticks or control console). The ROV operator has visual feedback on display from cameras and data from sensors (e.g., depth sensor, compass) to steer the ROV by sending control forces to the thrusters. The steering also includes a heading and depth control system that automatically keeps a given reference heading and depth provided by the operator. The raw data from the sensors are filtered to remove noises, such as inputs to the control system. The desired RPM (revolutions per minute) is allocated to each thruster, accordingly, from both the operator's and controller's input. The operator is provided with a rough heading and depth indication and the visual feedback to control the ROV. The control of the position and heading of the ROV to compensate for ROV dynamics and environmental disturbances (i.e. wind, waves and current) demands considerable effort from the operator. Moreover, preventing entanglement is dependent on the skill and experience of the ROV operator. The manual control mode is representative of LoA 1.

Station keeping mode and relative station keeping mode (LoA 2)

The ROV is equipped with a dynamic positioning (DP) system, to keep the position and orientation within certain excursion limits (Sørensen, 2012). For the ROV to hold its position and move along with the net, it depends on data fusion between various navigation sensors. The motion control system compensates for environmental disturbance and ROV dynamics. The navigation system is responsible for estimating the position, velocity, heading, depth, and altitude of an ROV in a given reference system. The navigation sensors need to make the system observable and controllable (Chen, 1998), which may contain a compass, a pressure gauge, and an ultrashort, short or long baseline network consisting of acoustic positioning systems (APSs), Doppler velocity log (DVL), gyroscopes and accelerometers (Dukan, 2014). The ROV senses any variation from the desired position, determines what thrust vector is required, and sends instructions to maintain the position and orientation.

In this operational mode, the ROV operator has the authority to take over control and control the thrusters directly. In this case, the operation switches modes from LoA2 to LoA1. Relative station keeping is similar to station-keeping, but instead of maintaining a constant position and heading, the ROV will maintain a constant distance and relative heading to the net. The initiation of relative station keeping requires the ROV to be close and pointing towards the net-pen.

Net-pen tracking mode (LoA 3)

The net-pen tracking is a desired semi-autonomous operation functionality (LoA 3) within aquaculture, currently under research (Duda et al., 2015; Rundtop and Frank, 2016). In this mode, the ROV follows the shape of the net to perform the inspection autonomously. Nonetheless, obtaining the relative position of the net pen is challenging, as the net-pen deforms by current-induced drag forces, which makes it an undetermined shape (Lader et al., 2008). For instance, a current velocity of 0.5 m/s may lead to a 20% volume reduction in an exposed net-pen (Lader et al., 2008). In this mode, DVL may function as a net relative sensor to provide net-pen relative velocities, range, and heading when the ROV is directed towards the net-pen. The desired velocities of the ROV can then be generated as the ROV moves along the varying net-pen, using the relative net-pen measurements. The guidance system of the ROV produces the desired reference velocity aiming to keep the ROV with a fixed net-pen relative range and heading, and keeping a plan for traversing the whole net-pen using the navigation system and the dynamics of the ROV. Note that the guidance system does not require prior knowledge of the whole net-pen.

In the net-pen tracking mode, the ROV operator is a supervisor, having the authority to intervene in case of an emergency or change of mission plan. This can be achieved by overriding (i.e., switch to station keeping mode or relative station keeping mode) or directly control the thrusters using joysticks (i.e. switch to manual control mode). In this mode, the operator provides a mission (i.e., an inspection of the whole net cage), with the ROV containing specific safe navigation rules in the presence of moving or static obstacles as an input to path re-planning. A set of waypoints are established according to the mission plan, the weather, the operation, among other factors. A smooth feasible trajectory is generated so that the ROV can follow. The raw data from the sensors, including images from the camera, are processed into the navigation system to give the relative net-pen position, position measurement, velocity, and heading estimates. Based on relative net-pen position and position measurement, a map is generated during operation to enhance the operator's operation monitoring. An example of such a representation of a fish cage is a 3D occupancy grid map (Hornung et al., 2013), which is sent to the graphical user interface (GUI).

Table 1
Selected System-level hazards and safety constraints.

System-level hazards	Safety constraints:
H-1: ROV unable to successfully complete the inspection	SC-1: ROV should be able to complete the inspection successfully by switching between operational modes properly SC-2: If ROV cannot complete the inspection, measures should be taken to ensure the task can be completed as soon as possible, especially in the situation when there is suspected structural damage or holes in the net that can lead to fish escape
H-2: ROV collides with net structure or tangles with the net	SC-3: ROV should not collide with obstacles under all operational modes SC-4: ROV should not tangle with the net under all operational modes SC-5: If the ROV collides or tangle with the net, the operator must acknowledge the situation and take measures to prevent making structure damage or holes in the net

Step 1.4 Describe tasks involving different operational mode - Net inspection task

To initiate the net inspection task, the operator launches the ROV into the fish cage. The mission is to traverse and inspect the entire fish cage. The ROV is initially set to manual control mode and manually manoeuvred to the side of the net. The operator configures the ROV for the preferred relative distance and heading to the net, as well as the desired velocity. Once the ROV is in position, the operator activates the net-pen tracking mode. The net inspection process is initiated, and the ROV operator supervises the operation and handles the ROV tether.

In case a hole is detected by the ROV, the relative station keeping mode is automatically activated, and an alarm is sent to the GUI to bring ROV operator's attention to the detected hole. The hole is examined and logged, and the ROV operator reactivates the net-pen tracking mode to continue the inspection. If the ROV is tangled in the ropes inside the

cage, the station-keeping mode is activated, and at the same time, the ROV operator is informed. The operator turns the ROV to manual mode to untangle itself. If the ROV is tangled in the net, the ROV operator is informed, and the thrusters are automatically turned off. A further investigation of the situation will then be needed to decide how to proceed. When obstacles are detected on the planned path (i.e., the produced desired trajectory), safe navigation rules are applied, and the ROV operator is informed. If the operator disagrees with the rule in certain situations, the ROV is switched to manual control mode, and the operator takes over to manoeuvre it to a safe location, and net-pen tracking mode is reactivated. Fig. 4 illustrates the main tasks performed during the net inspection and the operation modes involved:

3.2. Phase 2: model the control structures

Fig. 5 summarizes results from steps 2.1–2.5 in phase 2:

- Step 2.1 Model control structure
- Step 2.2 Assign responsibilities
- Step 2.3 Derive feedbacks
- Step 2.4 Repeat for each operational mode
- Step 2.5 Identify differences among modes

The control structures focus on two controllers: the human operator and ROV controller. ROV controllers need to collaborate to complete a mission by switching between operational modes at different LoAs upon different operational contexts. The lists of responsibilities specify the tasks each agent needs to accomplish for the safety constrains to be enforced. Due to limited space, Fig. 5 presents partial examples of assigned responsibilities at each operational mode for net inspection function and collision avoidance, with their corresponding process models and feedback. Note that the relative station keeping is not demonstrated separately due to its similarity with the station-keeping mode and net-pen tracking mode. The steps 2.1–2.5 are combined in one figure for the convenience of comparison among different operational modes.

3.3. Phase 3: identify unsafe control actions

Step 3.2 Identify operational mode transitions

The functional transitions between the four operational modes and shutdown mode are shown in Fig. 6.

Step 3.3 Identify triggering events

The triggering events for the ROV operational mode transition can be divided into two categories: external stimuli and internal stimuli, as defined in section 2.3. The triggering events are summarized in Fig. 7 and exemplified in Table 2.

External stimuli

When an ROV is on a mission, the human operator is an external stimulus who has the highest authority. The operator can always activate a transition of operational mode, following a call event (e.g., a request from an ROV operator). Another type of external stimulus is the change event, which covers the detection of interest (e.g., detect a hole in the net), a change in environmental features, and a detection of environmental objects.

The operational environment is critical to an ROV operation, and environmental interactions and consequent changes in environmental features should be anticipated. However, as pointed by Dogramadzi et al. (2014), the existing hazard identification methods do not encourage the safety analysis to consider different types of environmental interactions as an input to ensure safe robot operations. Changes in the environmental features are often associated with a change in sea conditions (e.g., sea state, a strong wind, strong current and tides, poor visibility due to fog or rain, salinity).

The detection of environmental obstacles includes the detection of

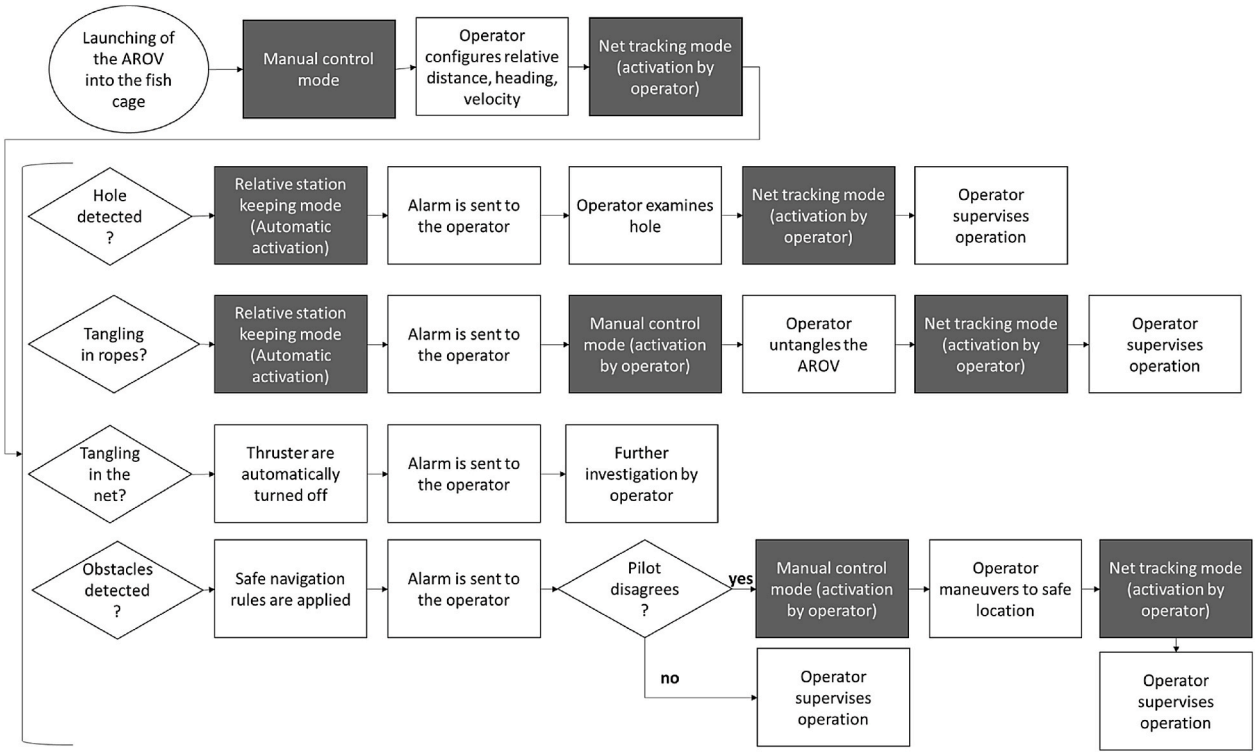

Fig. 4. Illustration of net inspection task and operation modes.

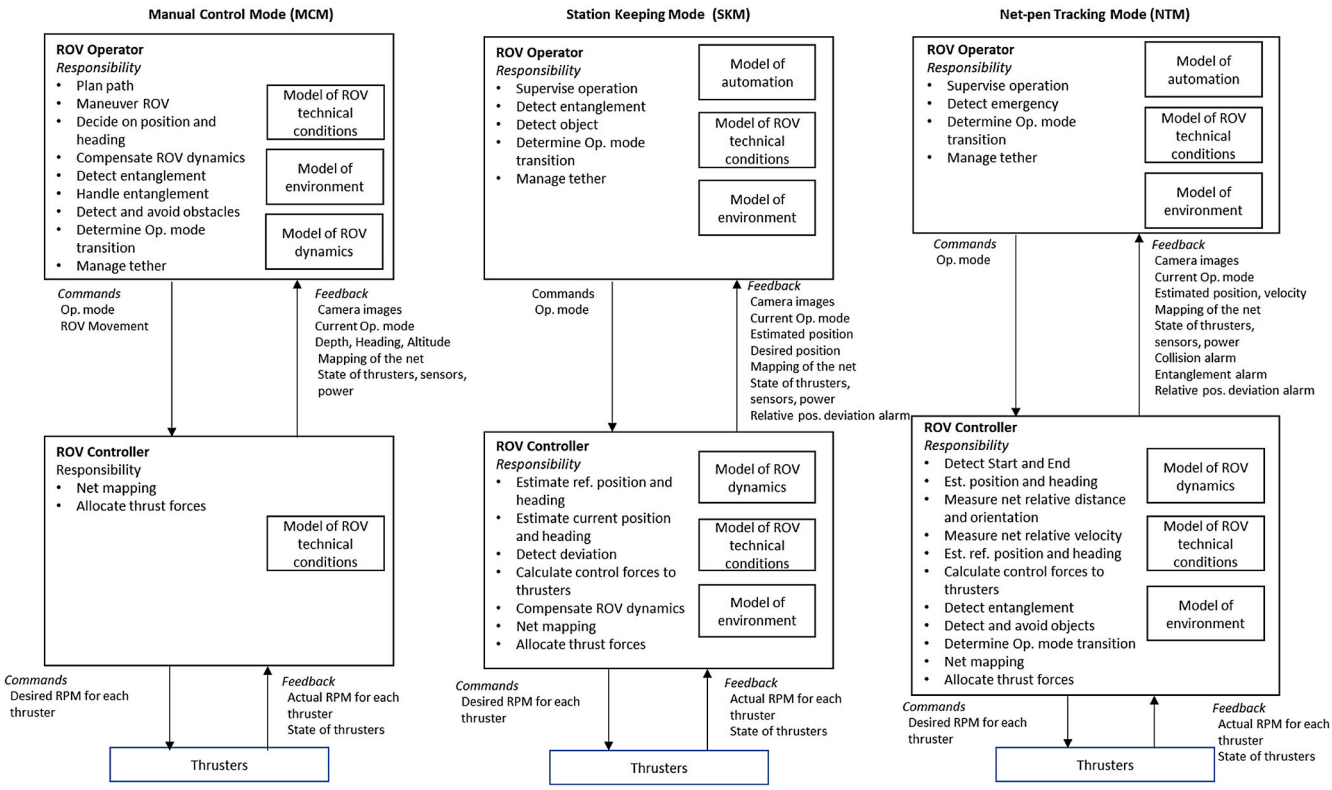

Fig. 5. Control Structures, responsibilities, feedbacks for ROV at different operational modes and LoA.

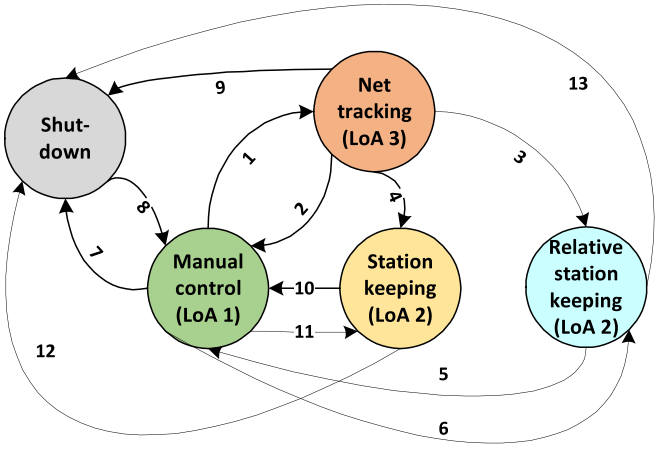

Fig. 6. Operational modes transition diagram.

objects and agents. Obstacles are related to obstructions to vehicle movements, such as fixed structures, surface floating obstructions (e.g., ships, buoys, anchor chains), objects suspended in the water column (e.g., fishing lines, loose netting) and bottom obstructions (e.g., subsurface structures, wrecks) (Christ and Wernli Sr, 2014). On the other hand, agents are the objects purposefully moving in the environment. Four categories of agents are suggested to capture the full range of behavioural patterns that any agent may exhibit and need to be perceived by the robot (Dogramadzi et al., 2014). They could be unintelligent (automatic systems), autonomous systems/other robots, animals and humans.

Internal stimuli

Internal stimuli take place within the ROV, which can cover timeout events, change events, and completion events. Timeout events may fire the transition when other awaited events do not occur within the specified time interval. The change events mainly originate from deviations of ROV performance, such as deviations in functional performance, technical failures, and software failures. An additional triggering event, or completion event, could also automatically initiate the transition from working mode to shutdown mode.

The ROV operator and ROV controller share the control and generate control actions following their responsibilities and update their process models based on feedback. Table 3 presents the responsibilities and process models that would enable a successful transition by a collaboration between the ROV operator and ROV controller. The process model is described by the status of process variables that would lead to a successful transition. Note that Table 3 explicitly describes the operator's responsibility of supervising the mission in cases where the operator's attention must be directed to one aspect of the operation, e.g., collision avoidance. Nonetheless, in addition to the ones identified in Table 3, the operator is responsible for the supervision of operations at all times and maintaining a situational awareness concerning the ROV mission and surrounding variables.

Table 4 presents the UTCAs for both the ROV operator and ROV controller for transition actions. The UTCAs are identified by considering the contexts for transition, responsibilities and process models described in Table 3.

3.4. Phase 4 identify loss scenarios

The next step is to identify scenarios and causal factors for each UCA, so that related safety constraints for safe transitions can be further detailed and refined. In Table 5, Table 6, Table 7, Table 8, some causal scenarios, causal factors, and related safety constraints for selected UCAs

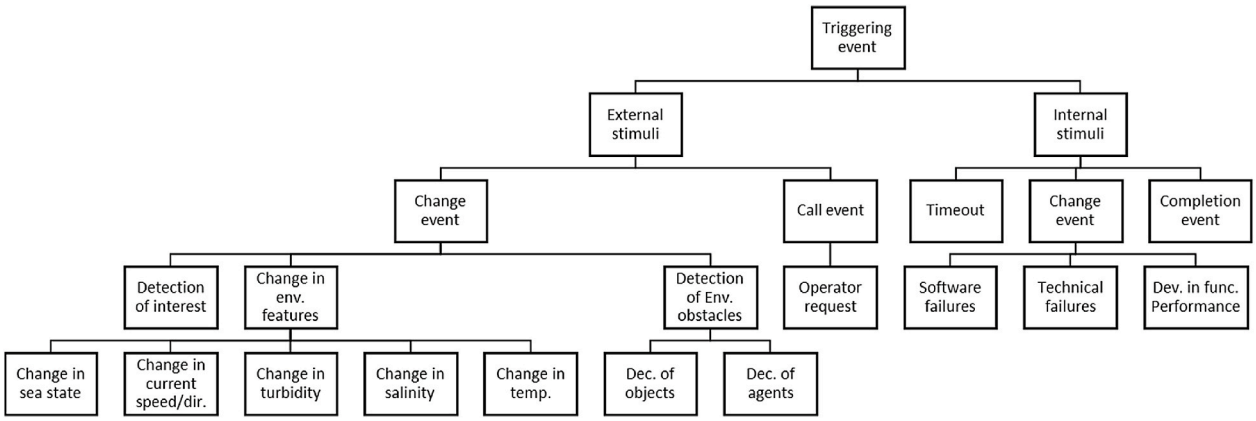

Fig. 7. Categories of triggering events for an autonomous system.

Table 2
Examples of triggering events, possible effects and possible transitions.

	Category	Sub-category	Examples of triggering events	Possible effects	Possible transitions (ref. Fig. 6)
External stimuli	Detection of interest	–	Holes in the net detected	–	3
	Change in environmental features	Change in sea state	Too high waves	Overstressed tether (>stress limit) due to heave, roll, pitch movement of the service vessel	2, 5, 10
		Change in current speed/direction	Too strong a current	ROV deviates significantly from relative distance and orientation	2, 5, 10
		Change in turbidity	Too high a turbidity	Degraded camera optics	4
		Change in salinity	Significantly decreased salinity	Acoustic navigation system failure	7, 9, 12, 13
		Change in temperature	Sharp temperature gradients	Thermal shock Acoustic navigation system failure	7, 9, 12, 13
	Detection of environmental obstacles	Detection of objects in collision courses	Detection of entanglement	Restricted movement of ROV	2, 5
		Detection of agents in collision courses	Detection of fish inside the cage	Blocked camera view by the fish	4, 11
	Operator request	–	Operator decides to further examine the net	–	2, 3, 4, 5, 10
Internal stimuli	Timeout	–	The ROV stays in one mode longer than a certain limit	ROV does not smoothly switch to the desired mode smoothly	7, 9, 12, 13
	Change event	Software failures	Collision avoidance failure in ROV controller	The ROV collides with obstacles	2
		Technical failures	One or several thrusters fail	The ROV controller must relocate forces to the rest thrusters, but the ROV may have limited manoeuvrability	2, 5, 10
		Deviation in functional performance	The ROV deviates from pre-defined relative distance	The ROV tangles with the net or misses the hole in the net	2, 9
	Completion event	–	The inspection of the whole cage is finished	–	7, 9

are presented.

4. Discussion

4.1. Findings from the case study

The ROV used in the case study is a representative autonomous marine system that has frequent operational mode transitions due to complex operating contexts inside of a confined area (i.e., inside of a fish cage). The ROV operator and ROV controller must collaborate closely to avoid tether entanglement, damage to the net structure, as well as inspect, report and temporarily block the holes (if necessary) in the net in time. The dynamic autonomy brought by shifting between various operational modes that have different LoAs adds an additional layer of complexity to ensure safe operation. However, this functionality has not been well addressed in current hazard identification methods. As commented on by Hollnagel (2005), co-agency specifically needs to be emphasized with a sound underlying model of the processes.

In this study, we proposed an approach based on STPA and a simplified state transition diagram in which a more explicit model of operational mode transitions is provided. The proposed method contributes to clearly communicating the shift of responsibilities by specifying i) how the allocation of the responsibility between human operators and the ROV controller changes under each operational mode; ii) what updated process model of the operator and the controller are to ensure a successful transition. The proposed approach makes it more visible as to what changes occur in the system when the operating mode is changed. The clarification of responsibilities under various situations can prepare the ROV operator better during the operation. For example, in NTM, the ROV needs to make decisions by itself to switch to RSKM or SKM and bring the ROV operator's attention to deal with the situation. The ROV operator is required to observe the deviations of functional

Table 3
Responsibilities and process models to enable successful transitions under specific contexts (selected results for each transition).

Transition No.	Triggering events	Context	Activated by	Responsibilities of the ROV operator	Process model of operator (i.e., operator's belief)	Responsibilities of ROV controller	Process model of ROV controller (i.e., ROV controller's belief)
1MCM-NTM	Operator request	Start autonomous net inspection after handling abnormal situation	ROV operator	Decide about abnormal situation Activate NTM button	Current mode = MCM Exceed current limits = NO Clear camera image = YES Collision danger = NO Send CMD = NTM Next mode = NTM	Receive commands Activate NTM function	Current mode = MCM Received CMD = NTM Next mode = NTM
2NTM-MCM	Collision avoidance software failure	ROV is on a collision course with certain objects in NTM	ROV operator	Supervise and assess the collision avoidance process Decide to take over control when integrated collision avoidance cannot handle the situation (e.g., due to software failure) Activate MCM	Current mode = NTM Collision danger = YES Next mode = MCM Send CMD = MCM	Receive MCM command Hand over control to ROV operator	Current mode = NTM If (collision danger is YES) then send alarm Received CMD = MCM Next mode = MCM
2NTM-MCM	Too strong a current	ROV deviates significantly from relative distance and orientation in NTM	ROV operator	Supervise and assess the relative distance from the user interface Decide to take over control Activate MCM	Current mode = NTM Relative position to net >Predefined relative distance to net = YES Next mode = MCM Send CMD = MCM	Monitoring current speed Measure relative distance and orientation Receive MCM command Handover control to ROV operator	Current mode = NTM Relative position to net >Predefined relative distance to net = YES Current speed > Limits Estimated next mode = NTM Received CMD = MCM Next mode = MCM

3NTM-RSKM	Holes in the net detected	A suspected hole is detected in NTM	ROV controller	Supervise and assess hole detection from the user interface
				Decide to take over control when ROV controller does not respond to a hole which is visible to the operator
				Activate RSKM manual when the hole is still in the vision of the camera
					Current mode = NTM
					Hole detected = YES
					Next mode = RSKM
						Switch to RSKM when ROV detects the hole
						Send notification to ROV operator
						Receive mode switch command when ROV does not detect the hole
							Current mode = NTM
							If (hole detected is YES)
							then set next mode to RSKM
							Else if (hole detected = No and Received CMD = RSKM),
							Then set Next mode to RSKM
							If (hole detected = No and no CMD received), then next mode = NTM

4NTM-SKM	Too high a turbidity	No clear net information is retrieved when turbidity is too high in NTM	ROV controller	Supervise from the user interface
				Take over control when ROV controller does not switch to SKM
				Activate SKM in time before ROV goes too far in such a context
					Current mode = NTM
					Clear camera image = NO
					Next mode = SKM
						Switch to SKM when no clear net information is retrieved
						Send notification to ROV operator
						Receive mode switch CMD when ROV continues moving under such a context
							Current mode = NTM
							If (clear camera image is NO)
							then set next mode = SKM
							if ((clear camera image = YES) and received CMD = SKM)
							Then set next mode = SKM
							if (clear camera image = YES and no CMD received), then next mode = NTM

5RSKM-MCM	Too strong a current	Too strong a current and the ROV cannot compensate for the dynamics on its own in RSKM	ROV operator	Observe ROV behaviour
				React upon alarm
				Decide to take over control
				Activate MCM
					Current mode = RSKM
					Exceed current limits = YES
					ROV behaviour = abnormal
					Send CMD = MCM
					Next mode = MCM
						Detection of too strong a current
						Inform ROV operator for interference
						Receive mode switch Command MCM
						Hand over control to ROV operator
							Current mode = RSKM
							if (exceeding current limits is detected), then send alarm to ROV operator
							if (exceeding current limits NOT detected) and (received CMD = MCM),

(*continued on next page*)

Table 3 (*continued*)

Transition No.	Triggering events	Context	Activated by	Responsibilities of the ROV operator	Process model of operator (i.e., operator's belief)	Responsibilities of ROV controller	Process model of ROV controller (i.e., ROV controller's belief)
							then set next mode = MCM
6MCM-RSKM	Operator request	The operator needs to examine the net further when the net is moving in MCM	ROV operator	Decide on further examination Activate RSKM	Current mode = MCM Send CMD = RSKM Next mode = RSKM	Receive RSKM command Activate RSKM function	Current mode = MCM Received CMD = RSKM Next mode = RSKM
7MCM-SD	Detection of objects	ROV operator detects ROV entanglement with the net in MCM	ROV operator	Detect entanglement Decide on the shutdown of the ROV	Current mode = MCM Net entanglement = YES Next mode = SD	Receive the shutdown command Send turnoff command to thrusters	Current mode = MCM Received CMD = SD Next mode = SD
8SD-MCM	Operator request	ROV starts operation in SD mode	ROV operator	Decide to start operation Activate MCM mode Manoeuvre thrusters via joysticks	Current mode = SD Next mode = MCM	Receive MCM command Send control forces to thrusters	Current mode = SD Received CMD = MCM Next mode = MCM
9NTM-SD	Detection of objects	ROV detects entanglement with the net in NTM	ROV controller	Supervise from the user interface Take over control when ROV controller does not switch to SD Activate SD in time before ROV damages the net	Current mode = NTM Net entanglement = YES Next mode = SD	Detection of entanglement with the net Send turnoff command to thrusters	Current mode = NTM Net entanglement = YES Estimated next mode = SD Next mode = SD
	Completion event	Inspection task is finished in NTM	ROV controller	Supervise from the user interface Take over control when ROV controller does not switch to SD Activate SD in time before ROV continues working for too long	Current mode = NTM Inspection finished = YES Next mode = SD	Detection of the ending of the net Send notification to ROV operator Send turnoff command to thrusters	Current mode = NTM Inspection finished = YES Estimated next mode = SD Next mode = SD

10SKM-MCM	Too strong a current	Too strong a current that ROV cannot compensate for with the dynamics on its own in SKM	ROV operator	Observe ROV behaviour
				React upon alarm
				Decide to take over control
				Activate MCM
				Manoeuvre thrusters via joysticks
					Current mode = SKM
					Exceed current limits = YES
					ROV behaviour = abnormal
					Send CMD = MCM
					Next mode = MCM
						Detection of too strong a current
						Inform ROV operator for interference
						Receive mode switch command MCM
						Hand over control to ROV operator
							Current mode = SKM
							If (exceeding current limits is detected), then send an alarm to ROV operator;
							if (exceeding current limits NOT detected) and (received CMD = MCM)
							then set next mode = MCM

11MCM-SKM	Operator request	The operator needs further to examine the net in calm sea state in MCM	ROV operator	Decide on switching to SKM Activate SKM	Current mode = MCM Next mode = SKM	Receive SKM command Activate SKM function	Current mode = MCM Received CMD = SKM Next mode = SKM
12SKM-SD	Detection of objects	ROV detects entanglement with the net in SKM	ROV controller	Supervise from user interface Take over control when ROV controller does not switch to SD Activate SD in time before ROV damages the net	Current mode = SKM Net entanglement = YES Next mode = SD	Detection of entanglement with the net Send turnoff command to thrusters	Current mode = SKM Net entanglement = YES Estimated next mode = SD Next mode = SD
13RSKM-SD	Detection of objects	ROV detects entanglement with the net	ROV controller	Supervise from user interface Take over control when ROV controller does not switch to SD Activate SD in time before ROV damages the net	Current mode = RSKM Net entanglement = YES Next mode = SD	Detection of entanglement with the net Send turnoff command to thrusters	Current mode = RSKM Net entanglement = YES Estimated next mode = SD Next mode = SD

performance and react promptly to take over (i.e., switch to MCM) or assist in switching mode. When the ROV entangles with the net and fails to turn off its thrusters, the ROV operator must detect the emergency as soon as possible and manually shut down the ROV. Another example is the situation that the turbidity is too high that the camera optics are significantly degraded. If the ROV does not switch to SKM automatically, the operator should manually activate the SKM in time before the ROV goes too far in the detection. The holes in the net might be missed.

The refined safety constraints based on the proposed approach also contribute to successfully completing a commission when there are dynamic control structures within one system. Some refined safety constraints from the case study especially emphasized the ensuring of safe transitions. Examples are as follows:

Table 4
Selected unsafe transition control actions for the operator and ROV controller.

Transition action	Context	Mode	Unsafe Transition Control Action (ROV operator)	Unsafe Transition Control Action (ROV controller)
Switch from NTM to MCM	ROV is on a collision course with an object in NTM (e.g., the safe navigation rules do not seem to work properly)	Not providing causes a hazard	[NTM-MCM]-UCA-OPERATOR-1: The ROV operator does not activate MCM when the ROV is close to colliding with an object in NTM [H-2]	[NTM-MCM]-UCA-ROV-1: The ROV controller does not send out a collision alarm when the ROV is close to colliding with an object [H-2]
		Provided causes a hazard	[NTM-MCM]-UCA- OPERATOR-2: The ROV operator switches to SKM/RSKM/SD rather than MCM [H-2]	[NTM-MCM]-UCA-ROV-2: The ROV controller switches to SKM/RSKM/SD, even though the MCM command is received [H-2]
		Provide too early/too late causes a hazard	[NTM-MCM]-UCA-OPERATOR-3: The ROV operator activates the manual control too late when an obstacle/agent is too close on a collision path [H-2]	[NTM-MCM]-UCA-ROV-3: The ROV controller sends out a collision alarm too late when an obstacle/agent is too close on a collision path [H-2]
	ROV deviates significantly from relative distance and orientation in NTM	Not providing causes a hazard	[NTM-MCM]-UCA-OPERATOR-4: The ROV operator does not activate MCM when the ROV deviates significantly from a predefined relative distance and orientation [H-1]	[NTM-MCM]-UCA-ROV-4: The ROV controller does not send a relative position deviation alarm when the ROV deviates significantly from a predefined relative distance and orientation [H-1]
		Provided causes a hazard	[NTM-MCM]-UCA-OPERATOR-5: The ROV operator switches to SKM/RSKM/SD rather than MCM when the ROV deviates significantly from a predefined relative distance and orientation [H-1, H-2]	[NTM-MCM]-UCA-ROV-5: The ROV controller switches to SKM/RSKM/SD, even though the MCM command is received [H-2]
		Provide too early/too late causes a hazard	[NTM-MCM]-UCA-OPERATOR-6: The ROV operator activates MCM too late, and the ROV deviates significantly from relative distance and orientation [H-1, H-2]	[NTM-MCM]-UCA-ROV-6: The ROV controller switches to MCM too late, when the ROV deviates significantly from its relative distance and orientation [H-1, H-2]
Switch from NTM to RSKM	A suspected hole is detected in NTM	Not providing causes a hazard	[NTM-RSKM]-UCA-OPERATOR-1: The ROV operator does not intervene when the ROV controller does not change mode [H-1]	[NTM-RSKM]-UCA-ROV-1: The ROV controller does not change mode when it detected a hole so the ROV continues moving down/up without logging the hole [H-1]
		Provided causes a hazard	[NTM-RSKM]-UCA-OPERATOR-2: The ROV operator intervenes and changes mode to MCM/SD, instead of RSKM [H-2]	[NTM-RSKM]-UCA-ROV-2: The ROV controller changes the mode to SKM, instead of RSKM, when detecting a hole [H-2]
		Provide too early/too late causes a hazard	[NTM-RSKM]-UCA-OPERATOR-3: The ROV operator intervenes too late when the ROV controller does not change mode in a timely manner so the ROV continues moving down/up without logging the hole [H-1]	[NTM-RSKM]-UCA-ROV-3: The ROV controller changes the mode too late when detecting a hole so the ROV continues moving down/up without logging the hole [H-1]
	The ROV is in the process of avoiding a collision with an obstacle	Not providing causes hazard	Not hazardous	Not hazardous
		Provided causes a hazard	[NTM-RSKM]-UCA-OPERATOR-4: The ROV operator switches from NTM to RSKM or SKM while an obstacle is approaching on a collision course [H-2]	[NTM-RSKM]-UCA-ROV-4: The ROV controller switches from NTM to RSKM or SKM while an obstacle is approaching on a collision course [H-2]
		Provide too early/too late causes a hazard	Not hazardous	Not hazardous
Switch from NTM to SKM	No clear net information is retrieved when the turbidity is too high in NTM	Not providing causes a hazard	[NTM-SKM]-UCA-OPERATOR-1: The ROV operator does not intervene when the ROV controller does not change the mode [H-1]	[NTM-SKM]-UCA-ROV-1: The ROV controller does not change modes when no clear net image is retrieved so the ROV continues moving down/up without logging the hole [H-1]
		Provided causes a hazard	[NTM-SKM]-UCA-OPERATOR-2: The ROV operator intervenes and changes the mode to MCM/SD, instead of SKM [H-1]	[NTM-SKM]-UCA-ROV-2: The ROV controller change the mode to RSKM, instead of SKM when no clear net image is retrieved [H-1]
		Provide too early/too late causes a hazard	[NTM-SKM]-UCA-OPERATOR-3: The ROV operator intervenes too late when the ROV controller does not change the mode promptly so the ROV continues moving down/up without inspecting some parts of the net [H-1]	[NTM-SKM]-UCA-ROV-3: The ROV controller change modes too late when the net image is not clear, so the ROV continues moving down/up without inspecting some parts of the net [H-1]

- The proposed collision avoidance path must be displayed in GUI to the ROV operator;
- The ROV operator must be notified when the operational mode switches;
- When the camera loses signals and does not respond to the ROV operator for a certain amount of time, the ROV must be shut down and floated to the surface;
- The ROV operator should be informed when objects enter dangerous zones, regardless of an evaluated (no) collision risk by the ROV controller;
- Training must be provided to the ROV operator to help become familiar with the capabilities of the system in various operational modes; and
- Another separate operator should be assigned for tether management.

The refined safety constraints are more concrete and can be input to improve system design, and operational procedures. When the higher-level autonomy is integrated into the system, the smooth transition faces the challenges of human action hazards such as skill degradation, loss of engagement, workload spikes, a lack of predictability associated with autonomous systems, and their overall complexity (Endsley, 2019). The human action hazards, in terms UTCAs for ROV operators, causal scenarios and the causal factors behind, need to be well considered while refining safety constraints. For example, it is clear from the results that the object avoidance functionality is not available in the SKM and RSKM modes. Without knowing the limitation, ROV operator may come to select the RSKM when an object is approaching on a collision course (i.e., how [NTM-RSKM]-UCA-OPERATOR-4 in Table 8 may occur).

In the case study, the proposed approach focuses on mode transitions and, as such, does not cover all aspects of the ROV. However, the analysis can be extended using standard STPA to analyse each operational mode and identify corresponding unsafe control actions and scenarios when the ROV is operated under such a mode. To limit the scope of this study, the above-mentioned analysis is not included in the case study.

Table 5
Causal scenarios and refined safety constraints for [NTM-MCM]-UCA-OPERATOR-1: The ROV operator does not activate MCM when the ROV is close to colliding with certain objects [H-2].

No.	Causal scenarios	Possible causal factors	Refined safety constraints
S1	The ROV operator does not realize the object is close to a collision	(a) The ROV controller did not send a collision alarm (b) The operator was informed but did not pay attention (c) The dead angle in the sensors	(a) Notification must be provided in time for the ROV operator when the minimum safe distance between the ROV and the object is violated (b) Design and test to ensure no dead angle in the sensors
S2	The ROV operator believes that the integrated collision avoidance function would be adequate to avoid a collision when an object is detected (but actually it is not)	(a) A lack of knowledge of the complexity of an autonomous system (e.g. software failures)	(a) The proposed collision avoidance path must be displayed in GUI to the ROV operator (b) Training of ROV operators so that they are familiar with the capabilities of the system and know what response to expect
S3	The ROV is close to the net, and the net is moving towards the ROV. The ROV operator is afraid to take manual control because a net collision may occur during the control transition.	(a) The operator would rather risk a collision with the obstacle than entangle with the net	(a) Notification must be provided in time for the ROV operator when the minimum safe distance between the ROV and the object is violated (b) the operator must acknowledge change of current direction
S4	The ROV operator detected the collision risk, but misjudged the movement of the object (that the agent will move in the opposite direction of the ROV)	(a) The operator has a lack of predictability associated with the collision avoidance function.	(a) Information about the type of the object (i.e. obstacle or agent) must be provided to the ROV operator.

4.2. Limitations and possibilities of proposed approach for autonomous operations

Investigating hazard identification for autonomous marine systems and operations, in general, can rapidly transform into a considerable effort. Nonetheless, it is a critical topic to be investigated. The proposed approach offers possibilities of application beyond the ROV case study presented in this paper. For a large part of possible applications of autonomy, autonomous systems with a LoA as high as “fully autonomous”, in which no human would be necessary or supervision tasks or remote control, is not expected in the near future. In this sense, these systems also have shared control. Many of them may also have possibilities of changing LoAs during the operations. An example is autonomous ships, which are also expected to have dynamic LoA. For instance, when unmooring out of a harbour with heavy traffic, the ship can be controlled by operators working onshore – remote control – and change the LoA to a higher level when reaching deep water with low traffic (Ramos et al., 2019b). Moreover, in case of a possible collision scenario, the operator may take over control of the ship, switching the LoA back to the remote control. Indeed, Ramos et al. (2020) states that the strong reliance on human-system interaction (shared control) and dynamic LoA present a challenge for risk assessments for autonomous ships.

The proposed approach allows for identifying triggering events for modes transition. This is particularly relevant for analysis of marine systems, which operate in a dynamic and complex environment. The existing hazard identification methods do not consider different types of environmental interaction systematically and sufficiently (Dogramadzi et al., 2014). The anticipation of possible environmental interactions is critical for a safe autonomous marine operation, particularly when the systems are moving toward the direction of being fully autonomous. The proposed method systematically identifies triggering events for transitions by dividing the events into external stimuli and internal stimuli events, and further into change event, timeout event, call event, and completion event. The results can serve as a basis to define triggering events in the design phase for other types of autonomous marine systems such as autonomous ships as well.

It is worth noting that, in some situations, simultaneous triggering events can create decision dilemmas. For example, when the ROV detects a suspected hole in the net, the ROV controller automatically switches to relative station keeping mode so that the hole can be logged and a further investigation can be carried out. If, at the same time, an agent is detected moving towards the ROV, the collision avoidance function is expected to be activated automatically as well. Under such a context, the best strategy might be to send a notification to the ROV operator and let the operator take over control to avoid a collision and then move back to a location to log information on the hole. If the hole is too big that fish start to escape, as an emergency procedure, the ROV should temporarily block the hole to mitigate fish escape until a remedy is implemented (e.g., send divers to repair). In such an emergent situation, the collision risk from moving the agent might be neglected. The identification of such decision dilemmas and the definition of corresponding strategies are critical to a robust and safe design of autonomous systems. The use of the proposed approach can shed light on identifying such dilemmas, through defined external and internal

Table 6
Causal scenarios and refined safety constraints for [NTM-MCM]-UCA-ROV-1: The ROV controller does not send out a collision alarm when the ROV is close to colliding with certain objects [H-2].

No.	Causal scenarios	Possible causal factors	Refined safety constraints
S1	The ROV controller does not detect an object in the collision path	(a) Camera failures (b) Navigation sensor failures (c) Too high an uncertainty in the estimated ROV position (d) Too high a turbidity (c) Dead angle in the sensors	(a) Signal processing module of the software must detect camera degradation and sensor failures and send alarms to the ROV operator (b) The software (e.g. the observer module of the ROV) must be designed to handle flawed or missing signals to the output's smooth estimated position and headings (c) When the turbidity is too high, the ROV controller should switch to SKM and inform the ROV operator (d) Design and test to ensure no dead angle in the sensors
S2	The ROV controller detected an object, but determined there was no risk of collision	(a) Failures in obstacle modeling (e.g. wrong judgement about the shape of the obstacle, or misjudging moving agents to obstacles) (b) Failures in risk evaluation in the integrated collision avoidance function	(a) Obstacle models must be tested by all different possible shapes and types of objects (incl. obstacles and agents) (b) The ROV operator should be informed when objects enter into dangerous zones, regardless of the evaluated collision risk by the ROV controller
S3	ROV controller started applying a safe navigation rule, but failed to avoid a collision	(a) Apply wrong rules to detected objects (e.g. apply rules for static objects to moving objects) (b) Failures in re-planning for a collision-free trajectory (c) Too strong of an environmental disturbance	(a) The type of object and detected moving path must be displayed in GUI (b) The ROV controller must evaluate whether the environmental load continuously exceeds operating limits

stimuli events (or conditions). This issue is not within the scope of this study but will be researched in further work. Furthermore, this study does not include the effects due to the interaction of external stimuli and internal stimuli (e.g., a stronger current and control system crash). Broader scenarios considering the possible interactions should be established in the later stage of STPA.

The literature presents STPA as a valuable method for hazard identification of autonomous and remotely controlled ships operation (Wróbel et al., 2018a, b), and the applicability to other autonomous marine systems needs to be tested in further work. The different characteristics, operating environment, sensing, motion, and reaction capabilities may bring new insights into the method. To this moment, STPA applications have not explicitly explored the hazards arising from control modes transition. The approach presented in this paper, which needs to be validated for other autonomous systems, can be an initial step towards the consideration of the important feature of control mode transition in STPA.

Table 7
Causal scenarios and refined safety constraints for [NTM-RSKM]-UCA-OPERATOR-3: The ROV operator intervenes too late when the ROV controller does not change the mode in a timely manner, so the ROV continues moving down/up without logging the hole [H-1].

No.	Causal scenarios	Possible causal factors	Refined safety constraints
S1	The signal delay between the image captured by camera and image displayed on displays	The communication cable blocks transmission from time to time	(a) Specialized connectors and cables must be used to maintain signal integrity at depth
S2	Camera signal loss so that the ROV operator does not get timely updated camera images	(a) The camera loses connection (b) Technical communication failure	(a) Camera signal loses connection alarm that must be sent to the ROV operator (b) When the camera loses signals, and there is no response from the ROV operator for a certain amount of time, the ROV must be shut down and floated to the surface
S3	The ROV operator does not realize the ROV controller does not change the mode promptly	(a) The operator is not informed about the situation promptly (b) The operator is informed in time but does not pay attention	(a) The ROV operator must be notified when the operational mode switches (b) The current operating mode must be displayed on displays
S4	The ROV operator does not interpret the situation correctly in time (e.g. the hole is soon going out of the vision of the camera)	(a) Lack of knowledge (b) Workload spikes (e.g. the tether needs to be handled at the same time) (c) The operator becomes distracted	(a) A separate operator should be assigned for tether management (b) Training should be provided about the working range of the camera

Table 8
Causal scenarios and refined safety constraints for [NTM-RSKM]-UCA-OPERATOR-4: The ROV operator switches from NTM to RSKM or SKM while an obstacle is approaching on a collision course [H-2].

No.	Causal scenarios	Possible causal factors	Refined safety constraints
S1	A hole is suspected by the ROV operator during NTM, but due to some turbidity, a further investigation is needed. The turbidity also causes the ROV operator not to detect an object on a collision course, and the ROV operator activates RSKM to investigate the suspected hole.	(a) Reduced visibility through the camera view	(a) The ROV must employ sensors, in addition to the camera, to detect objects on a collision course (b) The ROV controller must clearly inform the ROV operator when detecting an object on a collision course
S2	The ROV operator is aware that an object is approaching the ROV on a collision course, but believes that the ROV will perform collision avoidance in the RSKM	(a) The operator is not sufficiently familiar with the system	(a) Training to be conducted to ensure that operators are aware of the different system capabilities in different operating modes

5. Conclusion

This paper addresses the challenge of analysing the risk of autonomous marine systems, focusing on the first step of risk analysis, namely hazard identification. Autonomous systems may switch between different modes of operation and levels of autonomy, which increases the complexity, in particular with respect to interactions between the

human operator/supervisor, software, hardware, and the environmental conditions. These interactions impact risk but are challenging to identify and analyse. This aspect has not been paid enough attention in risk and safety studies of autonomous marine systems. The proposed approach in the paper uses STPA as a foundation, but explicitly includes dynamic control hierarchies representing the different LoA, and focuses on identifying unsafe transition control actions, scenarios, and causal factors. The proposed approach is exemplified for an ROV but is expected to be relevant for other autonomous marine systems such as autonomous ships. The results from the case study indicate that clarification of allocated responsibilities and updated process model of the operator and the controller are critical factors to ensure a safe operation. The defined safety constraints, as a result, focus specially on how to ensure the responsibilities and updated process models to be communicated clearly among controllers.

The proposed classification of triggering events shed light on the systematic identification of possible environmental interactions to improve the design of autonomous marine systems. Further work includes applying and testing the proposed method in a wide selection of domains, to further improve and develop a systematic approach to identify the triggering events and conditions for modes transition as input to eliminating hazards that may originate from unsafe transitions in the design phase.

CRediT authorship contribution statement

Xue Yang: Conceptualization, Methodology, Validation, Formal analysis, Writing - original draft. **Ingrid B. Utne:** Conceptualization, Writing - review & editing, Funding acquisition. **Stian S. Sandøy:** Validation, Writing - review & editing. **Marilia A. Ramos:** Visualization, Writing - review & editing. **Børge Rokseth:** Validation, Writing - review & editing.

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.

Acknowledgments

Yang, Sandøy and Utne's contributions to this paper have been carried out as part of the Reducing Risk in Aquaculture Project. The Norwegian Research Council is acknowledged as the main sponsor of project number 254913. Rokseth is funded by the Online Risk Management and Risk Control for Autonomous Ships (ORCAS) project. The Norwegian Research Council, DNV GL and Kongsberg Maritime are acknowledged as sponsors of project number 280655.

References

Bjelland, H.V., Føre, M., Lader, P., Kristiansen, D., Holmen, I.M., Fredheim, A., Grøtli, E. I., Fathi, D.E., Oppedal, F., Utne, I.B., 2015. Exposed Aquaculture in Norway, OCEANS'15 MTS. IEEE Washington. IEEE, pp. 1–10.
Burmeister, H.-C., Bruhn, W.C., Rødseth, Ø.J., Porathe, T., 2014. Can unmanned ships improve navigational safety?. In: Proceedings of the Transport Research Arena, TRA 2014, pp. 14–17. April 2014, Paris.
Chen, C.-T., 1998. Linear System Theory and Design. Oxford University Press, Inc.
Christ, R.D., Wernli, R.L., 2014. Part 2. The vehicle. In: Christ, R.D., Wernli, R.L. (Eds.), The ROV Manual, second ed. Butterworth-Heinemann, Oxford, pp. 53–54.
Christ, R.D., Wernli Sr, R.L., 2014. Chapter 3 - Design Theory and Standards, the ROV Manual, second ed. Butterworth-Heinemann, Oxford, pp. 55–92.
Dogramadzi, S., Giannaccini, M.E., Harper, C., Sobhani, M., Woodman, R., Choung, J., 2014. Environmental hazard analysis - a variant of preliminary hazard analysis for autonomous mobile robots. J. Intell. Rob. Syst. 76 (1), 73–117.
Duda, A., Schwendner, J., Stahl, A., Rundtop, P., 2015. Visual Pose Estimation for Autonomous Inspection of Fish Pens, OCEANS 2015 - Genova, pp. 1–6.
Dukan, F., 2014. ROV Motion Control Systems. Department of Marine Technology. Norwegian University of Science and Technology, Trondheim.
Endsley, M.R., 2019. The limits of highly autonomous vehicles: an uncertain future. Ergonomics 1–4.
Faisandier, A., 2013. Systems Architecture and Design. Sinergy'Com Belberaud, France.
Fan, C., Wróbel, K., Montewka, J., Gil, M., Wan, C., Zhang, D., 2020. A framework to identify factors influencing navigational risk for Maritime Autonomous Surface Ships. Ocean. Eng. 202, 107188.
Føre, H.M., Thorvaldsen, T., 2017. Causes for farmed salmon and trout escape during period 2010-2016. SINTEF Ocean, Trondheim.
Friedenthal, S., Moore, A., Steiner, R., 2014. A Practical Guide to SysML: the Systems Modeling Language, third ed. Morgan Kaufmann Publishers Inc, United States.
Harel, D., 1987. Statecharts: a visual formalism for complex systems. Sci. Comput. Program. 8 (3), 231–274.
Harris, C.A., Phillips, A.B., Dopico-Gonzalez, C., Brito, M.P., 2016. Risk and reliability modelling for multi-vehicle marine domains, Autonomous Underwater Vehicles 2016. AUV 286–293, 2016.
Heikkilä, E., Tuominen, R., Tiusanen, R., Montewka, J., Kujala, P., 2017. Safety qualification process for an autonomous ship prototype–a goal-based safety case approach. Marine Navigation 365–370.
Hinz, S.D., Hagenah, K.D., Pauli, H., 2010. Certification of unmanned underwater vehicles and working machines: safety and reliability under deep-sea and offshore conditions. IFAC Proceedings Volumes (IFAC-PapersOnline) 1–4.
Ho, G., Pavlovic, N., Arrabito, R., 2011. Human Factors Issues with Operating Unmanned Underwater Vehicles, Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage Publications Sage CA, Los Angeles, CA, pp. 429–433.
Holen, S.M., Yang, X., Utne, I.B., Haugen, S., 2019. Major accidents in Norwegian fish farming. Saf. Sci. 120, 32–43.
Hollnagel, E., 2005. Human reliability assessment in context. Nuclear Engineering and Technology 37 (2), 159–166.
Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W., 2013. OctoMap: an efficient probabilistic 3D mapping framework based on octrees. Aut. Robots 34 (3), 189–206.
Lader, P., Dempster, T., Fredheim, A., Jensen, Ø., 2008. Current induced net deformations in full-scale sea-cages for Atlantic salmon (Salmo salar). Aquacult. Eng. 38 (1), 52–65.
Leveson, N., 2011. Engineering a Safer World: Systems Thinking Applied to Safety. The MIT Press.
Leveson, N., Thomas, J., 2018. STPA Handbook.
Ludvigsen, M., Sørensen, A.J., 2016. Towards integrated autonomous underwater operations for ocean mapping and monitoring. Annu. Rev. Contr. 42, 145–157.
Norwegian Ministry of Trade Industry and Fisheries, 2017. Strategy on Escape Events (in Norwegian).
Ramos, M.A., Thieme, C.A., Utne, I.B., Mosleh, A., 2019a. White Paper, First International Workshop on Autonomous Systems Safety (IWASS), Trondheim, Norway.
Ramos, M.A., Thieme, C.A., Utne, I.B., Mosleh, A., 2020. Human-system concurrent task analysis for maritime autonomous surface ship operation and safety. Reliab. Eng. Syst. Saf. 195, 106697.
Ramos, M.A., Utne, I.B., Mosleh, A., 2019b. Collision avoidance on maritime autonomous surface ships: operators' tasks and human failure events. Saf. Sci. 116, 33–44.
Rausand, M., 2011. Risk Assessment: Theory, Methods, and Applications, 1 ed. Wiley, Hoboken, New Jersey.
Rødseth, Ø.J., Burmeister, H.-C., 2015. Risk assessment for an unmanned merchant ship. TransNav, International Journal on Marine Navigation and Safety Od Sea Transportation 9 (3).
Rokseth, B., Utne, I.B., Vinnem, J.E., 2017. A systems approach to risk analysis of maritime operations. In: Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 231, pp. 53–68, 1.
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.
Rundtop, P., Frank, K., 2016. Experimental evaluation of hydroacoustic instruments for ROV navigation along aquaculture net pens. Aquacult. Eng. 74, 143–156.
Schjølberg, I., Utne, I.B., 2015. Towards autonomy in ROV operations. IFAC-PapersOnLine 48 (2), 183–188.
Sintef Ocean, 2016. Centre for Research-Based Innovation on Exposed Aquaculture Operations.
Sørensen, A.J., 2012. Marine Control Systems Propulsion and Motion Control of Ships and Ocean Structures Lecture Notes.
Sørensen, A.J., Ludvigsen, M., 2015. Towards integrated autonomous underwater operations. IFAC-PapersOnLine 48 (2), 107–118.
Sperre Rov Technology, 2017. FLYING NET CLEANER, vol. 8.
Thieme, C.A., Utne, I.B., 2017. A risk model for autonomous marine systems and operation focusing on human-autonomy collaboration. In: Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, vol. 231, pp. 446–464, 4.
Thieme, C.A., Utne, I.B., Schjølberg, I., 2015. A risk management framework for unmanned underwater vehicles focusing on human and organizational factors. In: Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE.
Utne, I.B., Schjølberg, I., Holmen, I.M., Bar, E.M.S., 2017a. Risk Management in Aquaculture – Integrating Sustainability Perspectives, OMAE2017. Trondheim.
Utne, I.B., Sørensen, A.J., Schjølberg, I., 2017b. Risk Management of Autonomous Marine Systems and Operations, OMAE2017. Trondheim.

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. 190–202.
Valdez Banda, O.A., Kannos, S., Goerlandt, F., van Gelder, P.H.A.J.M., Bergström, M., Kujala, P., 2019. A systemic hazard analysis and management process for the concept design phase of an autonomous vessel. Reliab. Eng. Syst. Saf. 191, 106584.
Wróbel, K., Krata, P., Montewka, J., Hinz, T., 2016. Towards the development of a risk model for unmanned vessels design and operations. TransNav: International Journal on Marine Navigation and Safety of Sea Transportation 10.
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.
Wu, B., Cheng, T., Yip, T.L., Wang, Y., 2020. Fuzzy logic based dynamic decision-making system for intelligent navigation strategy within inland traffic separation schemes. Ocean. Eng. 197, 106909.
Xiang, X., Yu, C., Zhang, Q., 2017. On intelligent risk analysis and critical decision of underwater robotic vehicle. Ocean. Eng. 140, 453–465.
Xinqian, B., Chunhui, M., Zheping, Y., Jian, X., 2009. Reliability Analysis of AUV Based on Fuzzy Fault Tree. In: 2009 International Conference on Mechatronics and Automation, pp. 438–442.
Xu, H., Li, G., Liu, J., 2013. Reliability Analysis of an Autonomous Underwater Vehicle Using Fault Tree,. In: 2013 IEEE International Conference on Information and Automation (ICIA), pp. 1165–1170.
Yang, X., Utne, I.B., Holmen, I.M., 2020. Methodology for hazard identification in aquaculture operations (MHIAO). Saf. Sci. 121, 430–450.