Comparison of Hazardous Scenarios for Different Ship Autonomy Types using Systems-Theoretic Process Analysis

Hyungju KIM

Department of Marine Operations, University of South-eastern Norway (USN), Norway. E-mail: hyungju.kim@usn.no

Odd Ivar HAUGEN

Group Technology and Research, DNV-GL. E-mail: odd.ivar.haugen@dnvgl.com

Børge ROKSETH

Department of Marine Technology, Norwegian University of Science and Technology (NTNU), Norway. E-mail: borge.rokseth@ntnu.no

Mary Ann LUNDTEIGEN

Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Norway. E-mail: mary.a.lundteigen@ntnu.no

The area of autonomous and remotely operated ships is developing fast but is still an immature field where new ideas and novel technology solutions are being introduced. As part of these efforts, the Norwegian Forum for Autonomous Ships (NFAS) has defined six autonomy types for autonomous ships: two for continuously manned bridge, and four for fully or periodically unmanned systems. Different bridge manning levels and operational autonomy levels are allocated to each autonomy type, and therefore, each autonomy type could be associated with different kinds of scenarios leading to hazards. To support the decision making of the stakeholders, it is necessary to identify which autonomy type is related with which scenarios. The main objective of this paper is to identify and compare the scenarios leading to hazards of the six autonomy types. To analyse hazards of autonomous and remotely operated ships, we apply Systems-Theoretic Process Analysis (STPA). STPA is a relatively new hazard analysis technique that was developed to analyse hazards of modern complex and software-intensive control systems. STPA models the systems as a hierarchical control structure, and identifies scenarios leading to unsafe control actions that may lead to hazardous states or conditions. Six STPA analyses are conducted in this study to identify scenarios leading to hazards of the six autonomy types, and the results are compared and discussed.

Keywords: autonomous ship, unmanned ship, remotely operated ship, autonomy type, Systems Theoretic Process Analysis, STPA.

1. Introduction

The area of autonomous and remotely operated ships is developing fast but is still an immature field where new ideas and novel technology solutions are being introduced (DNV-GL 2018). As a part of these efforts, the Norwegian Forum for Autonomous Ships (NFAS 2017) defined six ship autonomy types combining four manning levels with three operational autonomy levels (NFAS 2017). Each ship autonomy type might be associated with different kinds of scenarios leading to hazards, because of different manning levels and operational autonomy levels. To ensure the safe operation of autonomous ships and to support decision making of stakeholders, it is required to identify and compare hazardous scenarios of each autonomy types.

We may need to apply a new hazard analysis technique for this, because the autonomous ships have more complex and software-intensive control systems than conventional human operated ships. These systems may lead to new types of accidents that are caused by unsafe and unintended interactions among the system components and controllers, rather than single component failures (Abdulkhaleq and Wagner 2013). The traditional hazard analysis methods may not effectively identify these new types of accidents (Leveson 2011). A recently developed hazard analysis technique, Systems-Theoretic Process Analysis (STPA), can identify hazardous scenarios that are not handled by traditional methods (Leveson and Thomas 2018).

The main objective of this paper is to identify and compare the scenarios leading to hazards of the six autonomy types, using STPA. For this purpose, we have investigated autonomous ships and ship autonomy types in Section 2, and STPA method in Section 3. In Section 4, we conducted STPA analyses for the six autonomy types, and the results are discussed in Section 5.

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio
Copyright © 2019 European Safety and Reliability Association.
Published by Research Publishing, Singapore.
ISBN: 978-981-11-2724-3; doi:10.3850/978-981-11-2724-3_0813-cd

2. Autonomous Ships and Ship Autonomy Types

2.1 *Autonomous Ships*

Accepting that autonomous ships are realized through higher degree of automation, we can look closer to what tasks need to be automated to realize an autonomous ship. Perhaps the first task that comes into mind is autonomous navigation - the systems' ability to navigate without human intervention. The navigation task includes functions like following a route, and anti-collision in compliance with COLREG (IMO Publications 2003). However, autonomous navigation requires that objects are *detected* and *classified*, and that their future spatial location in relation to the own ship can be predicted. The latter requires that, if the object is classified as another ship, its intensions and manoeuvrability must be predicted. For a fully autonomous ship, all these functions must be automatically and reliable be performed by the system. Some of these functions may also be performed by humans or in some combination between humans (onboard crew or remote control) and the autonomous system.

To be able to perform functions like following a prescribed route, or route (re-)planning, the current *and future* condition and capability (i.e. manoeuvrability) of the own ship must be known and predicted. This means that the machinery system’s power and thrust capability must be known, and that the system has the ability to alter its configuration to minimize the risk of performing the task at hand. Again, these functions may be performed by the system, humans, or in some combination.

Most, if not all functions mentioned so far, requires extensive numbers and types of sensors, many not historically used in the maritime industry, e.g. safe navigation requires reliable detection and classification of objects. That does not necessarily mean high-reliability sensors. Objects may be detected and classified though multiple sensors using different measuring principles (e.g. Radar, Lidar and an RGB camera may detect the same object). Therefore, the autonomous navigation system must be able to fuse the different sensors into a coherent world model.

These tasks and functions, partly or completely, may be performed remotely from a Shore Control Centre (SCC). This implies a number of new functions related to situation awareness (detection, analysis), means of engagement (act), and communication. Increased connectivity requires enhanced focus on cyber security.

2.2 *Autonomous Type*

The word autonomy is defined by Merriam-Webster (2019) as: "self-directing...", or "a self-governing". Other terms that comes into mind are: "free will", "independence", self-sufficiency", and perhaps "self-determination". With respect to technological systems, autonomy means self-directedness from us humans, i.e. there is no or limited human control over what the (control) system choses to do in a given situation.

In the industry there exists a number of definitions of autonomy targeting technological systems. These are based on level of autonomy, i.e. autonomy is not binary, but something in-between manual (dependent) and totally self-directedness (autonomous). Society of Automotive Engineers (SAE) defines six levels of autonomy going from 0 to 5 (SAE International 2016), and the International Association of Public Transport Union (UITP) defines four going from GoA1 to GoA4 (GoA: Grade of Automation) regarding train operations (UITP 2012). Underlying both definitions, lies the assumption that the degree of automation is closely related to the presence of the driver, and to what degree the driver is involved in the control of the vehicle/train. None of them include the possibility of remote control, i.e. low degree of self-directedness and no "driver" onboard the ship. In the maritime industry, remote control of ships is definitely a viable option. The definition of autonomy should therefore take into account the possibility that a ship may be controlled from a Shore Control Centre (SCC). In principal, a SCC can control more than one ship, and a ship may be controlled by more than one SCC, however, not simultaneously.

This paper uses the definition of Norwegian Forum for Autonomous Ships (NFAS) where the "level of autonomy" is defined along two dimensions (NFAS 2017): 1) bridge manning level, and 2) operational autonomy level. The operational autonomy level dimension is similar to other definitions of autonomy, but the "driver" dimension is removed from the definition. Different combination of the two dimensions defines *Ship autonomy types* as listed in Table 1.

Table 1. Two dimensions for ship autonomy type

	Manned bridge	Unmanned bridge - crew onboard	Unmanned bridge - no crew onboard
Decision support	Direct control no autonomy	Remote control	Remote control
Automatic	Automatic bridge	Automatic ship	Automatic ship
Constrained autonomous	-	Constrained autonomous	Constrained autonomous
Fully autonomous	-	-	Fully autonomous

Operational autonomy levels:

Decision support: The operation of the vessel is manual, and the crew gets decision support from advanced control systems using different sensors. This is the way most advanced ships are controlled today. Crew is in command and present on the bridge.

Automatic: Some well-defined advanced/demanding operations are achieved by the control system without human intervention. Contingency strategy involves human operators.

Constrained autonomous: The control system operates by itself in most situations within predefined limits. Humans must be available for immediate action if the control system cannot solve the situation. Humans may be located onboard or in SCC.

Fully autonomous: Ship handles all situations and operates in a fully self-governing manner.

Bridge manning levels:

Manned bridge: Today's situation

Unmanned bridge – crew on board: Crew may be mustered to the bridge to take control, but it will take some time.

Unmanned bridge: no crew onboard

Combining the two dimensions above creates ship autonomy types:

Direct control: Today's situation, crew is in control of the ship in all situation.

Automatic bridge: The crew monitors the control system and takes over control if necessary.

Remote control: Direct control from SCC.

Automatic ship: Monitored by SCC. SCC takes over if necessary.

Constrained autonomous: Monitored by SCC.

Fully autonomous: Human completely out of the loop.

3. Systems-Theoretic Process Analysis (STPA)

In Systems-Theoretic Accident Model and Processes (STAMP), loss of safety is seen as a control problem, and accidents can be prevented by means of successfully enforcing safety constraints through the application of control (Leveson 2011). Consequently, an accident occur due to inadequate control rather than as a consequence of failures. STPA is a method for identifying how such inadequate control may occur. This can, according to Leveson and Thomas (2018), be achieved in four steps.

In the first step, the purpose of the analysis is defined by specifying potential losses and the system level hazards that may cause these losses. The second step is to develop a hierarchical control model of the system. Note that the term controller is interpreted in a broad sense in STPA. Any entity exerting control over a process or another controller in the system is thought of as a controller. As such, computer-based control systems, human operators and organizations, such as the government (controlling both the design and operation of a system through legislation), can be viewed as controllers. Developing the control structure is achieved by identifying the relevant control loops and combining them into a control hierarchy.

The third step of STPA is to examine the control structure to identify potentially unsafe control actions (UCAs). Each control action is assessed to determine if and how it can lead to hazardous states by (i) not being provided, (ii) being provided, (iii) being provided too early or too late, or (iv) being applied for too long or too short.

The fourth and final step of STPA is to identify potential loss scenarios. These are scenarios in which the identified UCAs may occur. Such scenarios are typically identified by inspecting relevant parts of relevant control loops. Factors such as incorrect feedback, lack of feedback, control algorithm flaws, time-delay in actuators, component failures, disturbing process inputs and transmission errors in the control signal may be relevant to consider in the scenario development.

An advantage with STPA is that it can be applied at an early stage of the design and of a system and produce useful safety requirements and design constraints at this early stage. For example, in Rokseth, Haugen et al. (2018), as well as the present paper, STPA is applied to early concepts of autonomous ships. Another advantage of STPA is that it treats all system entities, such as mechanical subsystems, software, human operators and organizations, in the same way. This is possible because the analyst models the system as a control system. Any control entity is then treated simply as a controller rather than, for example, as a human operator or a computer control system (Rokseth 2018). In the application of STPA, the analysis team may face some challenges in the analysis, if the number of UCAs and loss scenarios gets very high. One approach is to apply a computerized tool, like XSTAMPP (Abdulkhaleq and Wagner 2015), to assist the analysis, or to apply a prioritization algorithm, as suggested in e.g. (Kim and Lundteigen 2018)

4. STPA to Six Ship Autonomy Types

In this section, we describe the approach and selected results from the STPA analysis that was conducted for the six autonomy types: (1) *direct control*, (2) *automatic bridge*, (3) *remote control*, (4) *automatic ship*, (5) *constrained autonomous*, and (6) *fully autonomous*. The analysis started by identifying and agreeing upon the main losses of concern: The resulting list was:

L1: Loss of life or injury to people (applies to manned ships only)
L2: Environmental pollution
L3: Damage to assets
L4: Financial losses due to unnecessary fuel consumption and/or delayed cargo

The analysis proceeded by identifying how the autonomous ship (regardless of the ship autonomy type) could, upon mal-functioning, lead to the identified losses. This resulted in a list of system-level hazards (MSC 83/21/2 2007) and system-level constraints to prevent these hazards, which are listed in Table 2. An important observation at this point of the analysis is that the results in Table 2 are not dependent on the ship autonomy type.

Table 2. System-level hazards and constraints

System-Level Hazard	System-Level Constraints
H-1: The ship strikes or is struck by another ship	SC-1: The ship must never strike nor be struck by another ship
H-2: The ship strikes or is struck by an external substance (but not a ship or the sea bottom)	SC-2: The ship must never strike nor be struck by an external substance
H-3: The ship touches the sea bottom	SC-3: The ship must never touch the sea bottom
H-4: The ship navigates in suboptimal speed and/or route	SC-4: The ship must always navigates in optimal speed and route

In step 2 of the STPA analysis, we had to prepare a control structure for each of the ship autonomy types, even if some parts and elements would be present in all. For example, all control structures consist (at the high level) of controllers (e.g., human operator, engine/steering control system and navigation system), actuators (main engines, steering gears), sensors (navigation sensors), physical system (hull), and disturbances. The control structures of the six autonomy types are provided from Fig. 1 to Fig. 6. The black arrows indicate control commands, blue arrows indicate feedbacks, and the green arrows indicate physical forces. The main distinctions between the six ship autonomy types are *where* the control is located, *how* controllers are realised, and *what types* of sensors that are needed to support the control. Based on our analysis, we propose the following six control structures, one for each of the six ship autonomy types as shown in Fig. 1 to Fig. 6. To ease the reading and comparison of the control structures, we identified in each illustration what elements are onshore (meaning remote from the ship) and which ones that are onboard.

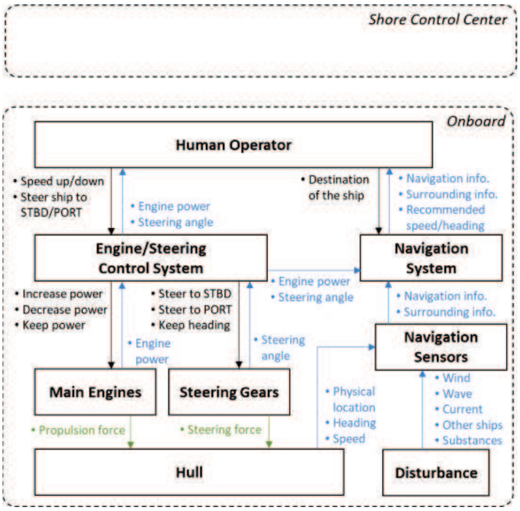

Fig. 1. Control structure of *direct control*

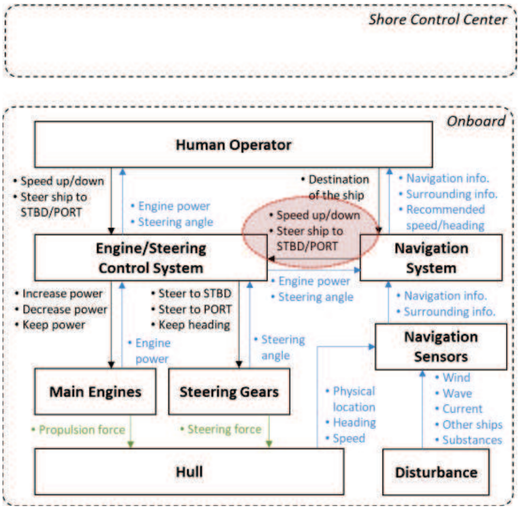

Fig. 2. Control structure of *automatic bridge*

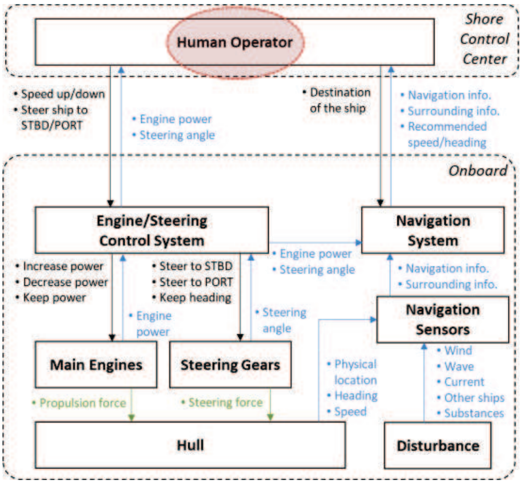

Fig. 3. Control structure of *remote control*

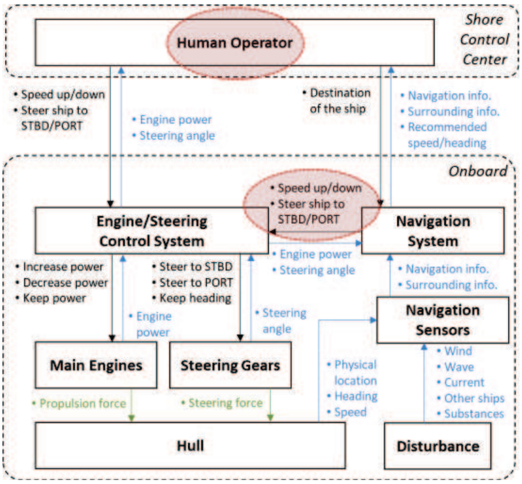

Fig. 4. Control structure of *automatic ship*

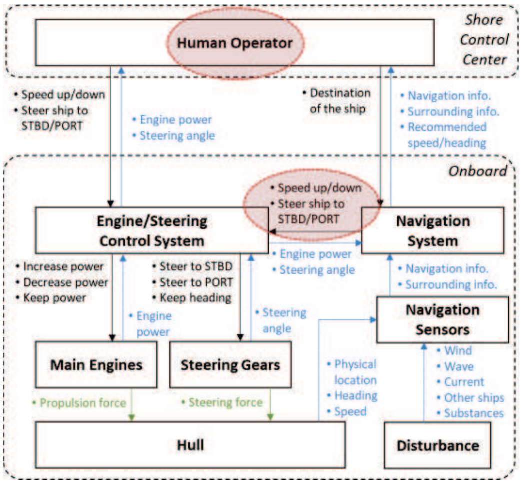

Fig. 5. Control structure of constrained autonomous

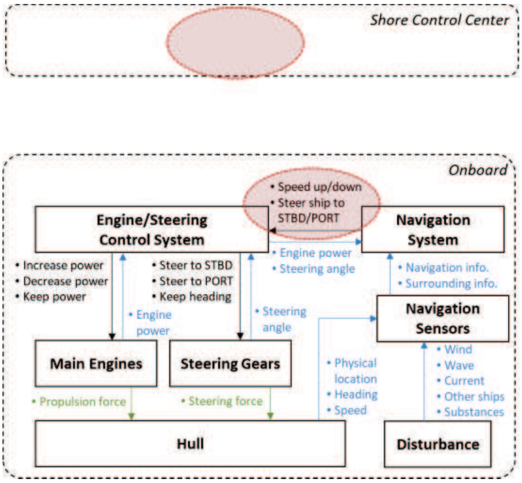

Fig. 6. Control structure of *fully autonomous*

Step 3 of the analysis focused on the identification of UCAs. For this paper, we decided to focus on illustrating how many UCAs we obtained for each of the ship autonomy type, and the distribution considering categories of controllers. Some examples of UCAs are provided in Table 3.

Table 3. Examples of UCAs of each autonomy type

Autonomy Type	Controller	UCAs
Direct control	Human operator on board	**UCA.HO.018** *Steer to STBD* command is not provided by human operator on board when it is needed to steer to STBD to avoid collision/ grounding/contact [H1, H2, H3]
Automatic bridge	Navigation system	**UCA.NS.018** *Steer to STBD* command is not provided by navigation system when it is needed to steer to STBD to avoid collision/ grounding/contact [H1, H2, H3]
Remote control	Human operator at shore control centre	**UCA.HO.002** *Speed up* command is not provided by human operator at shore control centre when it is needed to speed up the vessel to meet the voyage plan and it is safe to speed up [H4]
Automatic ship	Engine/ steering control system	**UCA.ES.022** *Steer to STBD* command is not provided by engine/steering control system to steering gears when this control command is originally provided by human operator at shore control centre [H1, H2, H3, H4]

Constrained autonomous	Engine/ steering control system	**UCA.ES.024** *Steer to STBD* command is provided too late by engine/ steering control system to steering gears when this control command is originally provided by navigation system [H1, H2, H3, H4]
Fully autonomous	Navigation system	**UCA.NS.004** *Speed up* command is provided too late by navigation system when it is needed to speed up to meet the voyage plan and it is safe to speed up [H4]

The results are presented in Fig. 7. From Fig. 7, it is observed that the two ship autonomy types *direct control* and *remote control* ended up with the same number UCA and distribution: 39 UCAs can be provided by the human operator and 42 UCAs by engine/steering control system. No UCAs were identified in relation to the navigation system. It is because the navigation system, for the two ship autonomy types, is used for decision-support only and no control command is provided by the system. For ship autonomy types *automatic bridge*, *automatic ship*, and *constrained autonomous*, we identified 74 UCAs from navigation system, in addition to the 81 UCAs of *direct control* and *remote control*. For the ship autonomy type *fully autonomous*, we identified 39 UCAs from navigation system and 42 UCAs from engine/steering control system, and there is no UCA from human operator (as expected).

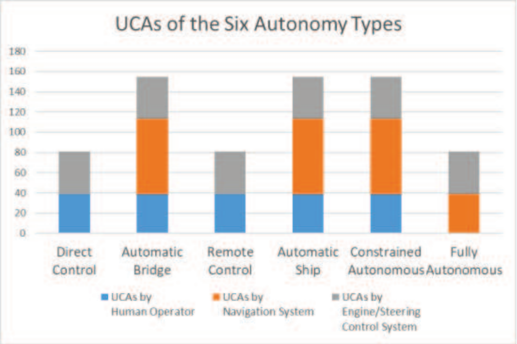

Fig. 7. Number of UCAs of the six autonomy types

From these UCAs, we developed the associated loss scenarios. The results were classified into five categories: technical failure, human error, control algorithm flaw, internal communication failure (inside the ship), and external communication failure (with shore control centre). Some examples of loss scenarios are listed in Table 4, and the results are illustrated in Fig 8.

Table 4. Examples of loss scenarios

Autonomy Type	**Loss scenario**
Direct control	**LS.HO.018.003** It is needed to steer to STBD to avoid collision/grounding/contact, but the human operator does not provide *steer to STBD* command because the human operator is not aware of this situation. This flawed process may occur if the navigation sensors fail, if power is not supplied to the navigation sensors, or if the signal cable from the navigation sensors it disconnected. As a result, the vessel may collide with a nearby ship and/or substance.
Automatic bridge	**LS.NS.018.004** It is needed to steer to STBD to avoid collision/grounding/contact, but the navigation system does not provide *steer to STBD* command because the navigation system incorrectly believes that the current route of the vessel is safe. This flawed process occurs if the navigation sensors provide wrong information about physical locations of nearby ships and/or other substances. As a result, the vessel may collide with a nearby ship and/or substance.
Remote control	**LS.HO.002.013** The human operator provides *speed up* command from shore control centre when it is needed to speed up the vessel to meet the voyage plan, but the vessel does not receive the control command due to communication failure between the shore control centre and the vessel. As a result, the vessel does not speed up and navigates behind the schedule.
Automatic ship	**LS.ES.022.007** The engine/steering control system provides *steer to STBD* command, and the steering gears correctly receive the command when this control command is originally provided by the human operator at shore control centre. However, the vessel does not steer to STBD due to no response from the steering gears. This flawed response may occur if the steering gears fail, if power is not supplied to the steering gears, or the steering gear is stuck. As a result, the vessel may collide with a nearby ship and/or substance.
Constrained autonomous	**LS.ES.024.005** The engine/steering control system provides *steer to STBD* command, and the steering gears correctly receive the command when this control command is originally provided by the navigation system. However, the vessel steers to STBD too late due to inadequate response from the steering gears. This flawed response may occur if the actuators of the steering gears fails to generate required physical forces.
Fully autonomous	**LS.NS.004.002** It is needed to speed up the vessel to meet

	the voyage plan, and every information is correctly received from the navigation sensors. However, inadequate software algorithm within the navigation system results in the *speed up* command being provided too late, and as a result, the vessel does not speed up on time and navigates behind the schedule.

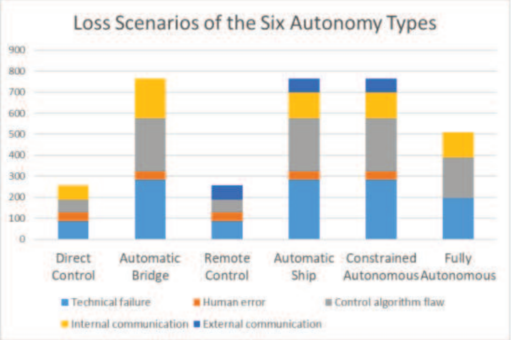

Fig. 8. Number of LSs of the six autonomy types

We observe in Fig 8 a similar trend as in Fig. 7. *Direct control* and *remote control* seem to share some common attributes and might be placed in one category (referred later as category A), and the same applies to the three types the *automatic bridge*, *automatic ship* and *constrained autonomous* (referred to as category B). *Fully autonomous* seems to be a category on its own (referred to as category C). This is elaborated further in the discussion.

5. Discussion

We observe from Fig. 7 and Fig 8 that the numbers of UCAs and loss scenarios are lower for category A, than for the other two categories. It seems therefore reasonable to ask: Does this mean that operation for category A is always safer? Our claim is that the answer is no. However, the magnitude of UCAs and the loss scenarios gives an indication of the complexity in the operation, which of course may have an impact on safety. To what extent safety is affected, depends on how well the system is able to handle the loss scenarios. What is interesting to note is that the number of UCAs (in Fig. 7) and the number of loss scenarios (in Fig. 8) are lower category C than for B. This might be a surprise at first glance, as the *fully autonomous* ship may be regarded as more complex than e.g. *automatic bridge*. However, another way to interpret this result is that the *mixture* of human interaction and automatic/autonomous systems can result in more situations of mal-operation than when the ship is *fully autonomous*. For example, according to Yang et al. (2018), hazards related to the timing of control hand-over are relevant in intermediate levels of automation but not in lower or higher level. This result can be seen as consistent with discussions seen in the automotive industry: The SAE level 3 “conditional automation” and SAE level 4 (high automation), which rely on the driver taking over in certain situations, might cause more dangerous situations than if the driver has full control. It is therefore discussed if autonomous cars, when implemented, should be *fully autonomous* (SAE level 5) rather than conditional. Even if not related, we see now in these days the discussions that are ongoing around the Boeing 737-Max 8/9. The plan has been grounded while the cause of two accidents (in Ethiopia and in Indonesia) are being investigated. One early hypothesis is that the autopilot has a fault, which rely on pilots taking over to compensate for an unwanted decline of plane nose (BBC 2019). No conclusion is made about the cause of accidents, but we can observe that this potential reliance on human interaction with autonomous systems might be more difficult to handle, in particular under stressful conditions.

6. Conclusion

This paper has discussed the application of STPA for vessels that apply different ship autonomy types. Some selected results have been and complemented with an elaborative discussion.

The main observation from the paper is that the combined reliance on humans and autonomous control actions *can* give a rise to more unsafe situations, than if the humans are in control or the ship is in full control. We have given some arguments why this might be a reasonable result. However, a more in-depth analysis with more detailed knowledge of actual implementations is needed to confirm its’ validity.

What seems reasonable to add as an observation is the reliance on the ability to capture all relevant loss scenarios. The loss scenarios are the basis for the specification of safety constraints, i.e. the measures needed to either avoid unsafe control or respond safely. An overlooked loss scenario may result in one or more missing safety constraints.

A last observation is that safety constraints have a very important role in the demonstration (verification and validation) of safety. The ability to capture the safety constraints into factory and site acceptance testing is therefore important, and further research may investigate how this might be done and how the test cases are selected.

References

Abdulkhaleq, A. and S. Wagner. (2013). Experiences with applying STPA to software-intensive systems in the automotive domain. In STAMP Conference at MIT.

Abdulkhaleq, A. and S. Wagner. (2015). XSTAMPP: An eXtensible STAMP platform as tool support for safety engineering. In STAMP Conference at MIT.

BBC (2019). Boeing 737 Max: What went wrong? Retrieved 23 April, 2019, from https://www.bbc.com/news/world-africa-47553174.

DNV-GL (2018). Class Guideline – Autonomous and remotely operated ships (DNVGL-CG-0264).

IMO Publications (2003). COLREG - Convention on the international regulations for preventing collisions at sea 1972, Consolidated Edition 2003. London, UK.

Kim, H. and M. Lundteigen (2018). Application of STPA to Subsea Systems - Opportunities and Challenges. Retrieved March 15, 2019, from http://psas.scripts.mit.edu/home/wp-content/uploads/2018/04/STPA-to-Subsea-20180329-Kim.pdf.

Leveson, N. G. (2011). *Engineering a safer world: Systems thinking applied to safety*. The MIT Press.

Leveson, N. G. and J. P. Thomas (2018). Cambridge, MA, USA.

Merriam-Webster (2019). Retrieved March 8, 2019, from https://www.merriam-webster.com/dictionary/autonomy.

MSC 83/21/2 (2007). Formal Safety Assessment - Container vessels. International Maritime Organization, London, UK.

NFAS (2017). Definition for Autonomous Merchant Ships. Norwegian Forum for Autonomous Ships, Trondheim, Norway.

Rokseth, B. (2018). *Safety and Verification of Advanced Maritime Vessels-An Approach Based on Systems Theory.* . Norwegian University of Science and Technology.

Rokseth, B., O. I. Haugen and I. B. Utne. (2018). Safety Verification for Autonomous Ships. In ICSC-ESWC 2018 Amsterdam.

SAE International (2016). Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. On-Road Automated Driving (ORAD) committee, PA, USA.

UITP (2012). Metro automation facts, figures and trends: a global bid for automation: UITP observatory of automated metros confirms sustained growth rates for the coming years. Retrieved 8th March, 2019, from https://www.uitp.org/metro-automation-facts-figures-and-trends.

Yang, X., Utne, I. B., & Thieme, C. A. (2018). A Review of Hazard Identification Techniques for Autonomous Operations in Norwegian Aquaculture. In Probabilistic Safety Assessment and Management PSAM 14. Los Angeles, CA.