The present and future of risk assessment of MASS: A literature review

Åsa Snilstveit Hoem
Department of Design, Norwegian University of Science and Technology, Norway. E-mail: aasa.hoem@ntnu.no

Internationally, there is an increasing interest in autonomous and unmanned ships, so-called Maritime Autonomous Surface Ships (MASS). This represents a paradigm shift that is presently underway promising safer, greener and more efficient ship traffic. A hypothesis of increased safety is often brought forward as we know from various studies that "human error" is the most frequently reported cause of marine casualties. In the latest Allianz report, the cost of losses resulting from "human error" between 2011 and 2016 is equivalent to 1.6 billion USD. Important questions in this context are; if we replace the human with automation, can we then reduce the number of accidents? And how can we evaluate the potential for new types of accidents to appear? The paper "At least as safe as manned shipping" by Porathe et al. (2018) presents a new risk picture and highlights the need for risk assessment. This paper continues on the risk assessment part by presenting a literature review of carried out in March 2018. More specific this paper gives a summary of five risk assessment methods presented in eight papers, and discuss their strengths and limitations, before addressing the main issues for future risk assessments of MASS.

Keywords: MASS(s), Unmanned ship, Literature review, Risk Analysis, Risk Assessment.

1. Introduction

Shipping is currently on its way into its fourth technical revolution, called Shipping 4.0 or cyber-shipping (Rødseth et al. 2016). Failure to anticipate and design for the new challenges that are certain to arise following periods of technology change can lead to automation surprises (Cook and Woods 1996). MASSs are no exception. MASSs may be low manned or unmanned (Rødseth et al. 2017). In principle, MASSs are required to be, at least, as safe as conventional surface ships in similar service (Jalonen 2017, Earthy and Lützhöft 2018, Porathe et al. 2018) To demonstrate a certain level of risk and evaluate if the safety goal is fulfilled, risk assessment should be carried out (Rausand 2013). A risk-based design approach is recommended to be used for the development of MASS by Lloyd's Register (2016) and DNV-GL (2018). One important question is then what is risk-based design, and what risk analysis should be carried out (in the design phase)?

2. Research questions

This paper addresses the following questions:
1. What risk identification analysis and methods for MASS can be found in the literature today? (Primarily an assessment of models of risk identification).
2. What are the main limitations and challenges of these risk assessments?

It is important to be concise in what is meant by MASSs, risks and risk analysis, in this paper's context. The next section presents the background and definitions used, followed by the method.

3. What is MASS?

The International Maritime Organization (IMO) currently use the term MASS for any vessel that fall under provisions of IMO instruments and which exhibits a level of automation that is currently not recognized under existing instruments. In the following, the term "autonomous ship" is a merchant ship that has some ability to operate independently of a human operator. This covers the whole specter from automated sensor integration, via decision support to computer-controlled decision making. An "unmanned ship" is a ship without crew that needs a certain degree of autonomy, e.g. to handle situations where communication with a remote shore control center (SCC) is lost.

Within Autonomous Marine Systems (AMS), underwater vehicles, especially Unmanned Underwater Vehicles (UUVs), have existed for several decades and are characterized through their capability to survey the subsea environment on a larger scale than divers and submarines are able to (Yuh et al. 2011). A taxonomy for the different types of autonomous maritime vehicles is proposed by the Norwegian Forum for Autonomous Ships (NFAS).

Typically, an autonomous system is a set of automated tasks, added interactions with several systems and/or human interaction, with capabilities and factors deciding the degree/level of autonomy. Frameworks for degrees or levels of automation (LOA) have been discussed by several professionals, mostly within the area of motor vehicle automation (SAE 2016, Vagia et al. 2016). Within the maritime domain, IMO has started a Regulatory Scoping exercise on MASS. Rigors discussions regarding definitions and characterization of ship autonomy are outside of the scope for this work.

MASS is a relatively new concept, mostly dating back to the MUNIN project (started in 2012), hence the classifications and proposed taxonomy are still evolving. Three main concepts are currently differentiated for MASSs:

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio
Copyright ©2019 by ESREL2019 Organizers. *Published by* Research Publishing, Singapore
ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0 esrel2019-paper

a. Low manned vessels with a partly unattended bridge (Bertram 2016, Rødseth 2017).
b. A swarm of MASSs supervised by one manned ship, so-called master-slave (Bertram 2016).
c. MASSs supervised from SCCs (Rødseth and Tjora 2014, Rødseth 2017).

A MASS with low manning (a.) is an intermediate solution to unmanned autonomous ship during the transition period (Bertram 2016).

4. What are risk assessment and risk analysis?

Risk is a term used in many contexts and in many different fields with different meanings. In general, risk also covers positive consequences, while in the majority of industries the focus is on negative consequences such as the risk of accidents (accident risk or security risk). Risk is the effect of uncertainty on objectives (ISO, 2018). It can be further defined as a combination of the potential events, their consequences and their likelihood. Risk assessment consists of risk identification, risk analysis and risk evaluation (ISO, 2018).

Risk analysis is the process to comprehend the nature of risk and to determine the level of risk (ISO, 2009). A source of danger that may cause harm to an asset is called a hazard (Rausand 2011). Reviewing hazards may identify sources of potential harm to the system, which gives input to a risk analysis. Component failure accidents have received the most attention in engineering, but component interaction accidents are becoming more common as the complexity of our system designs increases (Leveson 2012). The traditional view on risk assessment is to define the risk as the product of consequence and probability (Rausand 2003). For MASSs, a traditional risk analysis will attempt to find the likelihood of events, such as collision, allision, grounding, or stranding, and the assessment of consequences such as damage to people, the environment or to other ships or infrastructures. However, it should be noted that recent definitions of risk analysis take a broader, qualitative perspective to emphasize that not all uncertainties can be probabilistically expressed (Aven 2009). A common operational definition of risk analysis is the process of answering the following three questions given by Kaplan and Garrick (1981): 1. What can happen/go wrong? 2. How likely is it? 3. If it does happen, what are the consequences? These questions translate into three tasks:

- Hazard identification (examples are HazId, Hazard and Operability Analysis (HAZOP), Failure Modes, Effects, and Criticality Analysis (FMECA), System Theoretic Process Analysis (STPA) and blended hazard identification methodology)
- Causal analysis (like fault tree analysis (FTA))
- Consequence analysis (a barrier or exposure analysis, like event tree analysis (ETA))

As input to risk analysis, we use historical data inputs from similar operations (experience and learning from accidents) and knowledge about the system structure and design. However, historical data is non-existing for MASSs and knowledge about the system structure is limited, as the development of the first MASSs are still in a conceptual phase. Hence, it is of interest to see how the literature is addressing this lack of information, operational data and experiences with MASS.

5. Method

An initial literature search was conducted in December 2017 to establish a picture of what type of definitions are the most common ones, in the sense of number of results. As mentioned, autonomous vessels can be unattended, unmanned, and/or remotely controlled. When considering maritime safety, it is of interest to look for publications on potential accidents and

Table 1. Preliminary literature search

GOOGLE SCHOLAR [2011-2017]			Resultat	Relevant	SCOPUS [2011-2017]	Resultat	Relevant
Autonomous system safety (ship OR vessel OR ferry -underwater)	Unattended	Risk identification	16800	12	Risk identification	0	0
		Accident	15800	5	Risk	7	0
		Incident	15200	4	Accident or incident	5	0
	Unmanned	Risk identification	3570	10	Risk identification	13	1
		Accident	2150	10	Risk	111	3
		Incident	2400	14	Accident or incident	127	<5
	Remote contro	Risk identification	1800	8	Risk identification	1	0
		Accident	1140	6	Risk	32	<1
		Incident	1420	2	Accident or incident	23	<5
	Remotely contr	Risk identification	990	6	Risk identification	0	0
		Accident	556	4	Risk	11	<1
		Incident	612	4	Accident or incident	0	0

incident involving autonomous vessels, in the literature. Hence, several Boolean searches were carried out with strings of the following keywords and results in Google Scholar and Scopus:

5. 1 *Evaluation criteria for relevance*
To identify suitable and relevant publications, the following criteria had to be fulfilled: 1) The article must be related to the maritime domain, 2) From the title or abstract of the paper, the words "risk(s)" or "accident" and some level of automation must be present. The number of results in Google Scholar was overwhelming but the ratio of relevant literature versus the total number of results was quite low. After reviewing the title and abstract of the first 40 articles, the subsequent papers were not within the scope of the search and the review was limited to the first 50 articles of each string. Scopus, on the other hand, gave a lower, more manageable number of results. Nevertheless, one important finding is that "unmanned" got the most hits in the database of Scopus.

5.2 *Selection of papers*
A second literature review was conducted in March 2018. This time the literature was obtained through Boolean searches in three interdisciplinary databases; Scopus, Google Scholar and Web of Science. Based on the findings in the first study, "Unmanned" was selected together with the keyword "risk identification".

6. Relevant literature
The second literature search resulted in 42 documents. Most of the reviewed papers are articles published in scientific journals and papers presented at international conferences. From these, 18 papers were of interest.

7. Findings
Considering the timeline of the publications of interest, the results clearly indicate that the topic *autonomous* and *unmanned shipping* has increased in popularity in terms of publishing during the last decade. The authors are mainly researchers at Nordic Universities, the Netherlands, Poland, and Japan. Many of the papers link to the MUNIN-project and the AAWA-project presented in Jalonen et al. (2017).

In the literature search result, only eight of the papers concerns topics related to risk models and/or risk identification directly. From the literature, it is possible to see strong progress from 2013 towards risk models that could be useful today. As the eight papers present different approaches to the methods for risk identification, risk analysis, and risk management, each method (and paper) are listed separately in Table 2 below. Risk models are used to assess the risk arising from ship traffic or during ship operation. Goerland et al. (2015) reviewed the use of risk definition of published maritime risk models and concluded that in many cases the models do not state the risk definition or risk measure. This is also the case for the reviewed paper here. As insufficient data are available for MASSs, quantification of models is difficult and the risk models in the paper are of a qualitative nature. The models do not present a high level of detail in the model description or structure, hence making it difficult to assess and compare them. Hence, the next sections present and discusses each model or method for risk analyses separately.

7.1 *The MUNIN project's risk assessment framework (HazId, paper 1, 2 and 3)*
The MUNIN project developed a technical concept for the operation of an unmanned merchant vessel and assesses its technical, economic and legal feasibility. To be more specific, the core concept was a dry bulk carrier operating completely unmanned for parts of an intercontinental voyage. The concept relies on a SCC to handle complex situations. Analysis of collision and foundering scenarios for the concept concluded that a decrease of risk of around ten times compared to manned shipping is possible, mainly due to the elimination of crews' fatigue issues. The final report (Burmeister et al. 2014) states that risks of engine and other system breakdowns are expected to be lower for unmanned ships if proper redundancy is

Table 2. Overview of the relevant reviewed literature

No.	Author(s) (year)	Topic/Title	Risk model
1	Rødseth, Ø, & Tjora, A (2014)	A system architecture for an unmanned ship	HazId
2	Rødseth, Ø. & Burmeister, H.C. (2015)	Risk assessment for an unmanned merchant ship	
3	Rødseth, Ø. & Tjora, A (2015)	A risk based approach to the design of unmanned ship control systems	
4	Thieme, C.A. & Utne, I.B. (2017)	A risk model for autonomous marine systems and operation focusing on human–autonomy collaboration	BBN, HAC
5	Wróbel, K et al. (2017)	Towards the Development of a Risk Model for Unmanned Vessels Design and Operation	BBN, ETA
6	Utne, I.B. et al. (2017)	Risk Management of Autonomous Marine Systems and Operations	Risk management
7	Wróbel, K et al. (2017)	Towards the assessment of potential impact of unmanned vessels on maritime transportation safety	What If, HFACS
8	Wróbel, K et al. (2018)	System-theoretic approach to safety of remotely-controlled merchant vessel	STPA

implemented and improved maintenance and monitoring schemes are followed.

In 2013, Rødseth et al. published an "Unmanned ships operational context relationship diagram," and in 2015, a risk assessment framework was published. They present a risk-based structured approach to the design by controlling the risk elements while providing solutions for problems and document evidence that the risk level will be acceptable. The method presented here adopts parts of the Formal Safety Assessment method from IMO (2014). The initial architect structure (seen in Figure 1 below) is used as a basis in a HazId exercise to systemize the search for dangerous situations or risks.

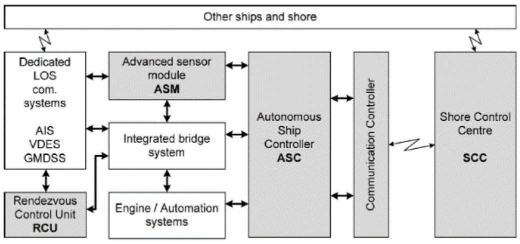

Fig.1: The MUNIN operational context relationship diagram derived from Rødseth et al. (2013).

In the risk assessment framework, different scenarios/accidents are considered, and hazards are identified together with mitigating actions, i.e. risk control options. They highlight the need to conduct a risk assessment before the system requirements are defined, in order to give input to the concept of operation (CONOPS) and verify the design. These risk control options aim at avoiding hazardous situations, but the interaction with the operator(s) is not given attention. Hence the method lack considerations of human autonomy collaboration. Given the paper is from 2015, the system architecture established here has formed the basis for other papers reviewed.

7.2 *A model describing the relationship between safety features of unmanned vessels (BBN, ETA, paper 5)*

With background in the MUNIN project and other sources where future anticipated design and performance are described (Burmeister et al., 2014; Rødseth and Burmeister, 2015), Wróbel, Krata, Montewka, and Hinz (2016) created a model describing the expected safety features in the paper "Towards the Development of a Risk Model for Unmanned Vessels Design and Operations". The risk model produced focuses on accidents' potential causes and failures within the system. The hazard analysis uses a Bayesian Belief Network (BBN) to describe the relationships between the safety issues from root cause to accident. The findings are structured into groups: Navigation, Engineering, Stability, and associated considerations, and Miscellaneous.

As the paper states, the model should be considered as a starting point to get an overview of relationships between safety features of unmanned vessels. There is no empirical data to support the likelihoods, so the validation is based on qualitative analysis. The model addresses several issues and potential accident types. However, addressing several accident types in one model may be a major challenge considering all different interactions and influencing factors. One major drawback is that the model does not include the communication connection to a SCC. In addition, the levels in the risk model are confusing; they are not levels of a technical system and should instead be considered as layers or steps or paths of an Event Tree Analysis (ETA). This assumption is made as the paper describe chains of consecutive events and conditions that may influence the consequences of potentially hazardous events. If the model were better structured, focusing on one accident type and explaining the levels and interactions, it would be useful as a basis for risk assessment of MASSs. Nevertheless, the paper addresses the challenge of uncertainties of the model due to unknown design and to the imperfection of brainstorming as a scientific method (Wróbel 2017, pp 7). From this publication in 2016, the model has been further developed.

7.3 *Review of marine accidents with "what if"-analysis and "HFACS" framework (paper 7)*

The same researchers (Wróbel et al., 2017) carried out a study of 100 marine accidents involving 119 vessels where the aim of the analysis was to assess whether the accident would have happened if the ship had been unmanned. It was also assessed whether its consequences would have been different. The assessment is based on a qualitative and subjective "what if"-analysis that ask: *1. If the ship were unmanned, how would that fact affect the likelihood of the particular accident?* and 2. *If the accident occurred anyway, would its consequences be more or less serious if there were no crew on-board?* The framework for Human Factors Analysis and Classification System for Marine Accidents (HFACS-MA) was set up to evaluate the causes of the accident. To answer the second question, the analysis of the accident's consequences was based on a simple check of whether the aftermath of maritime casualty affected people. The main challenge was the remoteness of the human operators, which has the benefit that the risk to the personnel is reduced. However, this remoteness implies that in case of an accident, like a fire, the human operator cannot recover the situation.

The "What-if" analysis and HFACS framework is not a method for risk assessment in the design phase, but the findings in the paper are of interest. It should be noted that the conditions for evaluating the safety of unmanned vessels are considering an unmanned vessel as a vessel where the bridge and crew is remote. The design and system architecture of autonomous systems might be completely different and new technology can cause accidents that we have not witnessed yet.. Another drawback of the study is the subjective evaluation of the effect of unmanned ships on the likelihood of the accidents and the many assumptions about which HFACS-MA causal category has the largest impact on an accident's occurrence. As a recommendation for further research the author emphasize the need to identify and list all anticipated hazards and their evaluated effects; only then can the level of safety associated with the unmanned ships operations be assessed (Wróbel et al., 2017, pp. 11).

7.4 *Systems-Theoretic Process Analysis, STPA (paper 8)*

In his latest paper, Wróbel argues that a framework building on the system-theoretic approach, STPA (Systems-Theoretic Process Analysis) is the best solution. This is also supported by Jalonen (2017). STPA is a hazard analysis technique based on STAMP (Systems-Theoretic Accident Model and Processes) first described by Leveson (2012). In order to perform a STPA, a safety control structure must be established. The safety control structure proposed in the paper is inspired by the many system architectures and models presented so far.

A list of hazards and correlated safety constraints related to different parts of the safety control structure is then presented. Furthermore, interaction mitigation of each control function is carried out in accordance with STPA principles. At the end of the paper (Wróbel et al., 2018) acknowledge the limitations of the system-theoretic method and present approaches on how to deal with uncertainties and so-called black swans. The modelling of the system is the most challenging part, causing a significant amount of uncertainty.

From the papers reviewed this is the most theoretically documented framework. However, from a safety perspective it could be beneficial to use a model that provides a quickly understood overview like the bow tie model that shows possible causal factors, consequences (outcome) and possible risk controls (barriers) linked to the hazardous events.

The analysis highlighted its preliminary status, addressing the uncertainty with respect to the design of MASSs. Technical issues have been identified as the factor contributing most to safety-related issues, followed by the interactions between SCC and the regulatory framework it needs to act under. Although interactions between operators are not covered. The authors of the paper try to add another dimension when including effectiveness and cost, which is not providing any useful information from a risk perspective. In an early design phase, knowledge of the system structure is limited. Hence, the mitigation of each interaction and their importance is valuable input to the evaluation and validation of design. Today, STPA are used for the assessment of dynamic positioning (DP) systems to identify hazards and for verification purposes (Rokseth et al., 2018).

7.5 *Bayesian Belief Network and Human Autonomy Collaboration (BBN, paper 5)*

As mentioned, Wróbel et al. (2016) suggest using Bayesian network for describing the relationships between the safety issues from root cause to accident. In a paper from 2017 Thieme and Utne investigate risk models focusing on human - autonomy collaboration. The main issue in the paper is that only a few risk models include human and organizational factors (HOFs). This aspect is illustrated in Figure 2 below.

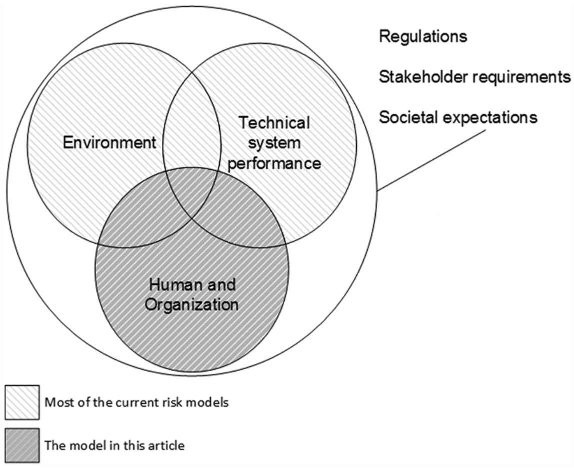

Fig. 2: The main aspects to include in an overall risk model. Derived from Thieme and Utne (2017).

The authors argue that risk models considering autonomous or remote operation should treat the human operators and the autonomous system as collaborators and not as individual or independent systems. The objective of the article is to present a BBN risk model focusing on human-Autonomy Collaboration (HAC) for AUV operation. Underwater vehicles are out of the scope of this paper. Nevertheless, as mentioned by the authors MASS may have similar requirements and demands as AUVs with respect to HAC, and the risk model could be adapted to other AMSs, as well (Thieme & Utne, 2017 pp. 1).

The paper provides a descriptive guideline for the steps involved in developing a BBN for risk modelling of HAC. This is a dynamic network of "nodes" which can be categorized as either Input-nodes, intermediate nodes and HAC nodes. The nodes have different states based on performance or status. The “arcs” connect the nodes (parent nodes to child nodes) and based on conditional probability tables (CPTs) for the parent node state, the child nodes’ state are determined. This way the BBN can be quantified and the human–autonomy collaboration performance can be assessed in order to identify relationships between technical, human, and organizational factors and their influences on mission risk. However, it is a wide-ranging task and data on the human operator performance is not easy to evaluate, as in the case of workload perception variability from operator to operator. Trust and overreliance are other ambiguous terms, which are influenced by several factors, which are not possible to model in BBN (Thieme & Utne, 2017).

This is a systemic accident model that sees accidents as a result of concurrent interactions at the system level, rather than individual failures. It can be considered as an alternative option to the STPA but could also as a supplement. STPA can identify the nodes and interconnections between operators, technical systems, and HOF. However, the advanced method is detailed and intricate, and requires an understanding of BBN that is not easily acquired. It should include all dynamic interactions of components and subsystems which is, as mentioned, an extensive task.

7.6 *Risk monitoring and control (paper 6)*

Utne et al. (2017) suggest a concept for risk monitoring and control for an autonomous ship. There is a clear distinction between risk assessment during the design phase and the operation phase in their work. This paper mentions HazId and BBN but most of the paper discourses the definition of risk and what risk assessment of MASS should include. Figure 3 below shows the proposed structure of a risk management framework.

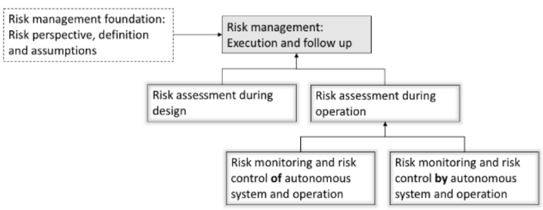

Fig.3: Risk management of autonomous marine systems, derived from Utne et al. (2017).

The other papers mainly concern operational risk, which only covers one lifecycle of the system/parts of the risk management. In this paper, the most common definition of risk, is presented with the added measure “strength of knowledge” and uncertainty. When the strength of knowledge is low, the uncertainty is high. The complexity of autonomous systems and operations is also highly related to uncertainty, as the more complex a system and operation is, the more difficult it is to gain “perfect” knowledge of it. This is a good point, but there is no sufficient way of including this measure in a risk analysis, other than subjectively considering assumptions, data quality and information available. The paper suggests using the BBN model presented in previous section. The paper recommends identifying hazards and risk influencing factors (RIFs) in the design phase and include uncertainty as the main constituent part of risk (rather than probability alone), nevertheless the paper do not go into depth on how to include this.

8. Human factors

The literature search did not include the term “human factors” or “human error”, however based on the findings in the literature listed in Table 2, this topic is given a section here. It is a consensus in the majority of the papers here that the contribution from human factors is important. Human factor issues and situation awareness are considered in five of the eight papers.

The explicit assumption is that with no humans on the bridge “human error” will go away (Porathe et al. 2018). The reason automation is safer is that they address human shortcomings like fatigue, limited attention span, information overload, normality bias etc. These issues are hypothesized to be reduced by increased ship autonomy by reducing the human involvement in direct control of ships, and by reducing the size of the crew on-board exposed to hazards of the hostile sea environment. However, it is important to remember that that our increasing dependence on information systems, and increasingly sharing of control of systems with automation, are creating a considerable potential for loss of information and control leading to new types of “human errors” (Leveson 2012).

There has been a cultural shift in the maritime industry toward increased levels of automation in tasks, particularly for navigation systems (Hetherington et al. 2006). This is partly because of reduced manning levels, as captains and crews are under increasing commercial pressure as supply chains are streamlined, and the availability of new technology. The paper “On Your Watch: Automation on the Bridge” by Lützhöft and Dekker (2002) discusses the qualitative consequences of automation on human work and safety. The paper propose that automation creates new human weaknesses and amplifies existing ones (Lützhöft and Dekker 2002 pp. 5). This is demonstrated by known accidents resulting from overreliance on machines. At the same time, automation can increase the cognitive demands on the reduced workforce.

In the discussion on “human error” it is important to remember that “human error” is not a cause but a result of other factors such as poor design, poor planning, poor procedures (Reason 2016). All human behavior is influenced by the context in which it occurs, and operators in high-tech systems are often at the mercy of the design of the automation they use. Hence, it might be more accurately to label an operator error as a flawed system or interface design instead. One example of this is a study of 27 collisions between attendant vessels and offshore facilities in the North Sea (Sandhåland et al. 2015). The study identified that errors due to reduced vigilance and misconceptions of the technical automation systems emerged as the primary antecedents of collisions.

Automation of human processes are expected to significantly reduce the number of incidents happening in shipping today. Nevertheless, the human element will not disappear. It will shift from ship to shore, where the remote operator exists and from where the software design and updating takes place. One must also assume that several potential accidents are adverted by the crew's actions and it is not clear if improved automation can match these numbers. Finally, one must also assume that some new types of incidents will occur because of the introduction of new technology and more automation.

9. More recent relevant literature

After the literature review was conducted several classification societies like DNV GL and Bureau Veritas (BV) have published guidelines on the topic of MASS and safety. DNV GL recommends the overall assurance process to be risk-based (DNV GL 2018), where minimum risk conditions (safe states) should be established based on structured risk analysis performed on several levels utilizing different methodologies; A preliminary hazard analysis (PHA) and a detailed risk analysis (FTA, ETA or FMEA), in addition to risk analysis method focusing on human aspects for operations from a SCC. BV also recommends assessing already available techniques for risk assessment (Veritas 2017).

10. Conclusions

It seems to be generally accepted that automation has the potential to decrease accidents that are due to human variability. However, automation has the potential of creating accidents, e.g. through transitions between automatic and manual control and the human having to rapidly assess the situation and make the right decisions. In the literature reviewed it seems that this challenge is not seemed to be included further than addressing situation awareness and human-machine (or autonomy) interaction.

Autonomy will create new types of accidents, partly due to accidents that was before averted by the human crew and partly due to introduction of new technology and corresponding new accident types. These types of accidents are challenging to include in the risk analysis as we lack statistical evidence for their probability. For further work it could be a good idea to make a database of the identified hazards and risks, and relate these to the dimensions of autonomy.

From the eight papers reviewed, it is difficult to conclude on one recommended practice for risk assessment of MASS. They all cover different topics, and some can be seen as overlapping and to some extent supplement each other. The papers highlight only parts of a socio-technical system, and a few scenarios. Some of the papers goes into depth in a case, while other papers highlight some perspectives and assumptions regarding the importance of safe operation and implementation. All risk analyses and models have different implications for how to analyze causes and consequences and target efforts. Comparing and discussing the results is hence challenging.

All eight papers acknowledge the lack of data on design solutions and system architectures and recognize that more work is necessary to develop approaches for risk analysis and assessment. Although, the STPA-method seems to be the most theoretically documented framework, it requires a high level of knowledge of the system architecture, and with a lack of empirical data subjective assumptions will be made to a greater extent. It should be stressed that all risk assessments and analysis have limitations. They also have different purposes and should be carried out both during the design and operation. Dynamic risk analysis will be important during operation, while risk assessment in the early design phase shall provide basis for constraints for the system, as pointed out by Utne et al. (2017). In the design phase it is beneficial to carry out a HazId/PHA and iterate it with the CONOPS until all relevant risks are managed.

As mentioned, no empirical studies have been performed to compare and evaluate the reviewed methods for risk analyses of MASSs. According to a study by Thieme et al. (2018), risk assessment and modelling of AUV and Autonomous Remotely operated vehicles are presented with operational data to some extent in the literature (Thieme et al. 2018 pp. 12-13). While this is not the case for MASS where less research has been conducted, both on the qualitative and quantitative side of risk analyses, as of today.

11. Recommendation for further research

From the review, the following main challenges, and hence request for further research within risk assessment in the design of MASSs, is listed:

- The need to cope with the lack of empirical (historical) data

- The need to include the human operator in the loop. In highly automated and autonomous systems, the influence of operators and other people interacting with the system is unneglectable (Bainbridge, 1983)
- The need for improved causal models to explicitly model organizational factors and software failures
- The need to consider dependencies between systems, including safety and security issues in complex control actions of MASS

Acknowledgement

This work is supported by the SAREPTA (Safety, autonomy, remote control and operations of industrial transport systems) project, financed by the Norwegian Research Council with Grant No. 267860.

References

Aven, T., & Renn, O. (2009). On risk defined as an event where the outcome is uncertain. Journal of risk research, 12(1), 1-11.

Bainbridge, L. (1983). Ironies of automation. Analysis, Design and Evaluation of Man–Machine Systems 1982, Elsevier: 129-135.

Burmeister, H. C., Bruhn, W., Rødseth, Ø. J., and Porathe, T. (2014). Autonomous unmanned merchant vessel and its contribution towards e-Navigation implementation: The MUNIN perspective. *International Journal of e-Navigation and Maritime Economy, 1, 1-13*.

Cook, R. and Woods, D. (1996). Adapting to new technology in the operating room. *Human factors, 38(4), 593-613.*

DNV-GL (2018). DNV GL Class Guideline 0264 Available:http://rules.dnvgl.com/docs/pdf/dnvgl/cg/2018-09/dnvgl-cg-0264.pdf [Accessed 31.10.2018]

Earthy, J. V. and Lützhöft, M. (2018). Autonomous ships, ICT and safety management. In Managing Maritime Safety (pp. 143-165).

Hetherington, C., et al. (2006). "Safety in shipping: The human element." *Journal of safety research 37(4).*

IMO, I. M. O. (2007). Guidelines for the application of Formal Safety Assessment (FSA) for use in the IMO rule-making process. MSC/Circ.1180-MEPC/Circ.

Jalonen, R. T., et al. (2017). Safety of Unmanned Ships: Safe Shipping with Autonomous and Remote Controlled Ship. Aalto University.

Kaplan, S. and Garrick, B. J. (1981). On the quantitative definition of risk. Risk analysis, 1(1), 11-27.

Leveson, N. (2012). Engineering a safer world: systems thinking applied to safety. Cambridge, MIT Press.

Lloyd's Register (2016, February), "Cyber-enabled ships. Deploying information and communications technology in shipping". First edition.

Lützhöft, M. H. and S. W. A. Dekker (2002). "On Your Watch: Automation on the Bridge." *Journal of Navigation 55(1).*

Porathe, T., et al. (2018). At least as safe as manned shipping? Autonomous shipping, safety and "human error". Safety and Reliability–Safe Societies in a Changing World, CRC Press: 417-425.

Rausand, M. (2013). Risk assessment: theory, methods, and applications, John Wiley & Sons.

Reason, J. (2016). Managing the risks of organizational accidents, Routledge.

Rokseth, B., Utne, I. B., & Vinnem, J. E. (2018). Deriving verification objectives and scenarios for maritime systems using the systems-theoretic process analysis. *Reliability Engineering & System Safety, 169*, 18-31.

Rødseth, Ø. and Tjora, Å. (2014). "A risk based approach to the design of unmanned ship control systems." Maritime-Port Technology and Development 2014.

Rødseth, Ø. J. and Burmeister, H. C. (2015). Risk assessment for an unmanned merchant ship. *International Journal on Marine Navigation and Safety Od Sea Transportation,* 9.

Rødseth, Ø. J., et al. (2016). "Big data in shipping-Challenges and opportunities."

Rødseth, Ø. J., et al. (2017). "Characterization of autonomy in merchant ships."

SAE (2016). "SAE J3016: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles."

Sandhåland, H., et al. (2015). "Situation awareness in bridge operations–A study of collisions between attendant vessels and offshore facilities in the North Sea." *Safety Science 79: 277-285.*

Thieme, C. A. and Utne, I.B (2017). "A risk model for autonomous marine systems and operation focusing on human–autonomy collaboration." *Journal of Risk and Reliability 231(4): 446-464.*

Thieme, C. A., Utne, I. B., and Haugen, S. (2018). Assessing ship risk model applicability to Marine Autonomous Surface Ships. *Ocean Engineering.*

Utne, I. B., et al. (2017). Risk Management of Autonomous Marine Systems and Operations. ASME 2017 *36th International Conference on Ocean, Offshore and Arctic Engineering.*

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?. *Applied ergonomics,* 53, 190-202.

Veritas, B. (2017). Guidelines for autonomous shipping. Guidance Note NI, 641.

Woods, D., et al. (2017). Behind human error, CRC Press.

Wróbel, K., et al. (2017). "Towards the assessment of potential impact of unmanned vessels on maritime transportation safety." *Reliability Engineering & System Safety* 165: 155-169.

Wróbel, K., et al (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*.

Yuh, J., Marani, G. and Blidberg, D. R. (2011). Applications of marine robotic vehicles. *Intelligent service robotics, 4(4), 221.*

Proceedings of the 29th European Safety and Reliability Conference.
Edited by Michael Beer and Enrico Zio
Copyright ©2019 by ESREL2019 Organizers. *Published by* Research Publishing, Singapore
ISBN: 981-973-0000-00-0 :: doi: 10.3850/981-973-0000-00-0 esrel2019-paper