Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P090-cd

A Criticism of Proposed Levels of Autonomy for MASS

Ørnulf Jan Rødseth

Norwegian Forum for Autonomous Ships, Norway. E-mail: ojr@mits-consult.com

Lars Andreas Lien Wennersberg

SINTEF Ocean, Norway. E-mail: lars.andreas.wennersberg@sintef.no

Levels of Autonomy (LOA) is a popular subject in both scientific and regulatory literature. This applies to many different types of autonomous systems, but here we will focus on "Maritime Autonomous Surface Ship" (MASS). LOA is often looked at as a scale on which one can measure autonomy, as if it is a gradually emerging property of a system. We will argue that it is more useful to look at autonomy as a binary concept, it is either present or it is not. There are differences in what functions the autonomy covers and under what conditions autonomy can be used safely, but that is not something that can easily be measured on a numeric scale.

Keywords: Levels of Autonomy, Operational Envelope, Constrained Autonomy, MASS

1. Introduction

Levels of Autonomy (LOA) is a popular subject both in scientific and regulatory literature (Vagia, Transeth, and Fjerdingen 2016; Rødseth, Wennersberg, and Nordahl 2022b). This applies to many different types of autonomous systems, but this paper focuses on Maritime Autonomous Surface Ships (MASS), i.e., ships above 500 gross tons that may sail on international voyages.

In the maritime area, LOA is often used to describe how automated or how independent the operation of a MASS is from human supervision or intervention. LOA may also cover if crew resides on the ship or are located elsewhere. Different schemas have been proposed to capture the various combinations.

The International Maritime Organization (IMO) used four degrees of autonomy during their regulatory scoping exercise for MASS (IMO 2018) that covers four combinations of automation levels and where operators reside.

Although not directly relevant for MASS, The Central Commission for the Navigation of the Rhine (CCNR) has proposed six levels of automation for inland vessels (CCNR 2022). The structure of these six levels are based on the structure of the six levels of driving automation as proposed by the Society for Automotive Engineers (SAE), for application to road vehicles (SAE 2021). This classification focuses mainly on what functions automations takes care of, i.e., longitudinal versus transversal control, and how fallbacks, i.e., when automation is no longer capable of control, are handled.

Lloyd’s Register (LR) proposes seven autonomy levels from AL0 to AL6 (Lloyd’s Register 2017) while DNV defines five levels for navigation functions (DNV 2021). These classifications defines the relationship between operator and automation, e.g., via decision support and operator on the loop, to full automatic control without operator intervention.

Sheridan and Verplank (1978) defined 10 levels of supervisory control that is often cited in relationship to LOA. However, these 10 levels are more describing different stages in a decision pipeline where the human is effectively in control up to level 6 and the automation takes over control from level 7 and up.

ALFUS (Huang et al. 2005) is a different approach where both degree of automation as well as environmental and mission complexity is taken into consideration. This is a form for multi-dimensional quantification of autonomy.

Another approach is taken by the International Organisation for Standardization (ISO) in its Vocabulary for Autonomous Ships (ISO 2022) . Here, an informative annex links different degrees of human control (DoC*)* and degrees of automation (DoA) together. The result is a more descriptive version of the relationship between the humans and automation systems for different of degrees of autonomy (DOA). We will come back to this in section 5.

It is also necessary to consider for what purpose a LOA is used. Some LOAs are used to do comparisons between different autonomous systems, such as ALFUS. Others are used to classify relationships between increased automation to existing rules and regulations, such as the IMO system. Others, like DNV and LR are used to define new requirements to systems with increasing levels of automation. Our interest is mainly to use the LOA to classify the respective "areas of responsibility" of automation and human when controlling the ship. This is directly related to defining necessary operational procedures for humans and corresponding requirements to automation.

Section 2 will provide our definitions of autonomy and automation and how the operator fits into that picture. Section 3 gives a brief note on the issue of operators being locally present or not., where it is argued that this should not be a part of the definition of autonomy. Section 4 discusses how this leads to the concepts of constrained autonomy, operational design domain and operational envelope. Section 4 discusses a paradigm where we have a real "cooperation" between humans and automation to keep the ship efficiently and safely sailing. This includes the definition of different operational modes and a discussion on the difference between having crew onboard or in a remote control centre (RCC). Section 6 will give a brief argument on why full autonomy is not so relevant for MASS. Section 8 discusses operator control modes as opposed to autonomous operation and section 8 gives a summary of the different processes that are needed on a ship and the implications for ship autonomy. Section 9 contains our conclusions.

2. Automation versus Autonomy

The difference between autonomy and automation has been discussed for a long time and will continue to be discussed (Rødseth and Vagia 2020; Rødseth, Wennersberg, and Nordahl 2022b). Although there are many suggestions as to how differences can be determined, it can also be argued that there is no discernible difference at all. This has led the Society of Automotive Engineers (SAE 2021) and the Rhine Commission (CCNR 2022) to depreciate the term autonomy and use respectively "driving automation" and "automated navigation" instead.

However, as argued in our previous papers, we believe that a distinction is useful and that a useful definition can be provided for both terms. These definitions have been published in ISO/TS 23860 (ISO 2022) and, slightly transcribed, they read as follows:

Automatic: Processes or equipment that, under specified conditions, <u>can</u> function without human control.

Autonomous: Processes or equipment that, under certain conditions, <u>are designed and verified to be controlled by automation</u>, without human assistance.

These definitions establish that automation is doing the same thing in both cases, but that autonomy appears when the automation can be *trusted* to control processes or equipment under the specified conditions. This does not completely avoid the ambiguity as shown below by examples related to a ship autopilot:

(i) The autopilot is *autonomous* with respect to control of the ship's safe sailing when there are no obstacles ahead. Automation can be trusted to control the ship if the operator knows that there is no chance of collision for a sufficient time ahead.

(ii) The autopilot is *automatic* with respect to control of the ship's safe sailing in areas with obstacles. An operator needs to continuously assess if it is still safe to sail on autopilot and take control if it is not. Automation cannot be trusted for the full scope of operations.

(iii) The autopilot is *autonomous* with respect to controlling the ship's safe sailing in areas with obstacles if it is connected to an anti-collision radar that can alert an operator sufficient time before a dangerous situation can occur. The autopilot can be trusted to sail the ship until the alert is activated.

The autopilot is doing the same automated function, but the property of autonomy is dependent on if it can be trusted to perform the function without human supervision or assistance.

This is captured in the phrase "*under certain conditions*" in the definitions. Thus, we will argue that autonomy is a property that either exists or do not exist. Autonomy cannot as such be graded in levels.

Ships are complex systems with many different processes, where navigation is only one. In addition, one will need energy production, stability management, fire protection and more. The term "*processes or equipment*" captures that different processes may have different degrees of automation. As an example, fire protection may need to be autonomous under the whole voyage, while navigation may only be autonomous under conditions where anti-collision systems can be trusted to detect problems in time for operators to assist.

3. Remote control or crew onboard

Some LOA, e.g., the IMO definition, include the presence of crew onboard in the classification. As we argue in (Rødseth, Wennersberg, and Nordahl 2022b), this does not change the definition of autonomy, it only changes the maximum response time that operators need for intervention if automation finds out that it is no longer able to provide autonomous control. The issue of response times will be discussed in the next section and remote versus local presence will not be discussed further in this paper.

4. Constrained Autonomy

The previous section argued that autonomy requires that automation can be trusted to control a given process under certain conditions. When these conditions change, the operator needs to be alerted to handle the situation. However, this raises the question of how much time the operator needs to be able to safely take over control. This question is not easy to answer as it will depend heavily on the means to gain situational awareness and will likely be different for people located in an RCC than for those on a ship's bridge. In Lu, Coster and de Winter (2017), experiments are done on car driving that indicates that seven seconds may be sufficient to safely take over control. The process to gain situational awareness on a ship or in a RCC is different from that in a car (Yoshida et al. 2020), and we can also assume that the situational displays or tools will be different from what we see on a ship's bridge today (Ottesen 2014). Thus, it is not possible here to say exactly what time is required, but somewhere less than a minute for an RCC operator, from getting an alert that action is needed until he or she can safely take over control is in our opinion very likely, given the above referenced studies. One can here assume a control room setup as that suggested in the MUNIN project (Porathe 2014), with each operator being responsible for a maximum of six ships and having back-office operators for handling more complex incidents that the front-end operator cannot immediately correct.

The issue of "out-of-the-loop loss of situational awareness" (Endsley 2017) is an obvious addition to the problem. We should avoid having one operator continuously monitoring one ship, where most of the time nothing happens. A better approach is to let the automation system alert the operator when attention is needed, freeing the operator to do other things when automation is in control. However, this requires that the automation can detect that conditions change so that operator attendance may be needed, early enough for the operator to gain situational awareness and safely take corrective actions.

We have in a previous paper (Rødseth, Wennersberg, and Nordahl 2022a) introduced the maximum response time or T_{MR} for the time the operator needs to get sufficient situational awareness to control the ship safely. Likewise, we defined the minimum deadline or T_{DL} as the minimum time the automation system can "look ahead" and reliably predict that the operator will not be needed to take over control. These time constants can be used to define a necessary criterion for safe use of autonomy by requiring that $T_{DL} > T_{MR}$ under the prevailing operational conditions.

In the car industry, (SAE 2021) defines the Operational Design Domain (ODD) as *"Operating conditions under which a given driving automation system or feature thereof is specifically designed to function."* This implies that operator attention is needed immediately after the system exits the ODD. To satisfy the above timing criteria, the automation must limit itself to a subset of the ODD where it is still time for the operator to regain control when conditions change. We propose the term "constrained autonomy" for this form of autonomy, as the autonomy is "self-constrained" to this subset. Likewise, we can define the "Constrained Autonomy Domain" (CAD) as the conditions under which the automation system can safely control a ship process and reliably alert the operator in time to safely take over control when these conditions change.

Most autonomous ship projects known today, e.g. Yara Birkeland and ASKO (Felski and Zwolak 2020), have restrictions on where they can operate. Currently, this is mainly related to a defined geographic area in national and relatively sheltered waters. This means that there are limits to what conditions the ship can operate under, even when humans assist the automation. The Operational Envelope (OE) is proposed as a term for the combined capabilities of the automation systems and the humans that are involved in MASS operations, and is by ISO (ISO 2022) described as the "*conditions and related operator control modes under which an autonomous ship system is designed to operate, including all tolerable events.*"

This also means that these transitions between operator and automation control is a designed feature of the system and are not considered to be fallbacks. The ISO standard also defines fallback as a "*designed state that can be entered through a fallback function when it is not possible for the autonomous ship system to stay within the operational envelope*".

Note that the standard uses the term "fallback" rather than "minimal risk condition" or similar. This is because the requirement is that a fallback state should *not result in an intolerable risk,* and this may be either in the "as low as reasonably practicable," "acceptable" or "negligible" risk regions (Melchers 2001). It is in general not a *minimum* risk.

Thus, the different operational domains can be illustrated as sub-sets of each other as seen in Fig. 1. The fallback space consists of the defined fallback states and will be outside the operational envelope.

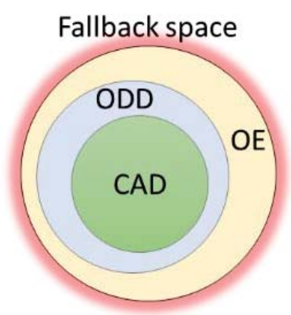

Fig. 1. Nested operational domains.

5. Operator and Automation Cooperates

The assignment of trust to the automation system to enable autonomy, means that the operation of the autonomous ship can be seen as a true cooperation between human and automation. Only one of them is in control at a given time and when automation is in control, within the CAD, the operator is free to do other things than continuously overseeing the automation.

This defines three cooperation modes where the relationship between T_{DL} and T_{MR} is also defined:

- Operator exclusive (OE): Outside ODD where operator may be assisted by the automation but is in control and needs to be immediately able to intervene ($T_{DL} = 0$, $T_{MR} = 0$).
- Operator Assisted (OA): In the ODD but outside CAD. The automation can still control the ship but needs continuous attention by the operator. The operator is in control but can use own judgement to leave the control position for shorter periods ($T_{DL} \geq 0$, $T_{MR} = 0$).
- Constrained Autonomous (CA): In the CAD where automation is safely in control and will alert operator before intervention is needed. Automation is in control ($T_{DL} > T_{MR}$).

When the system is in ODD, but outside CAD it means that there is a $T_{DL} \geq 0$, but that its value is not known by the automation, hence, continuous operator attention is needed.

In previous papers, we have also suggested a fourth mode: Full Autonomy (FA, $T_{DL} = \infty$), where no operator is needed at any time during the ship's operation. As we will come back to in section 6, this is an unlikely mode for MASS, but can be used to give a more complete overview of all operational modes.

As stated in section 3, the difference between having crew onboard or in an RCC is the maximum response time (T_{MR}). During working hours with crew on or near the bridge, response times would be similar to that of RCC operators. However, if autonomy was to be used to allow crew to go to sleep during night-time oceanic passages and with no RCC crew as backup, the maximum response time would be significantly longer, e.g. 20 minutes. This would not change the operational modes but would require that the automation system must have a correspondingly longer minimum deadline (T_{DL}), i.e. 20 minutes.

This may be feasible on open sea and fair weather, with a good long-range radar for anti-collision.

6. Full Autonomy or Not?

The concept of “full autonomy” or “full automation” is used in most proposals for levels of automation or autonomy. Most authors, including in IMO, agree that full autonomy means that the operator is or may be completely out of the control loop. However, it can be discussed whether this is a viable concept, at least for ships:

(i) Normally, one will want to use the ship for a specific purpose, which requires that one can instruct it to do the intended task. The task may also change during the ship's voyage, requiring an update to the instructions. This is not consistent with full autonomy.

(ii) There is a fair degree of agreement in the legal community that future interpretation of international legislation may allow that a ship’s Captain to be in the RCC (Ringbom and Veal 2017; Van Hooydonk 2014). There is even less doubt that you will need a Captain somewhere and that the Captain must have some measure of control over the ship. Thus, no full autonomy.

(iii) Ships are high value assets, and it is not likely that the owner will leave it to its own devices without some form of supervision and possibility for control.

(iv) There is also doubt that the ship can be safely designed to operate fully without human assistance within today's rules and legislation (Porathe et al. 2018). This also puts the idea of full autonomy into doubt, at least for the near future.

(v) Components will ultimately fail, or the ship will encounter conditions outside its operational envelope. In these cases, some form of fallback will be activated, but this will most likely require human intervention to restore the ship-functions to normal or to salvage the ship.

(vi) If there is indeed an RCC, even only for supervision, it does not make economic sense to develop automation that can handle all and every situation. As an operator is available, it is in most cases more cost-effective to leave some of the more complex situations to that operator (Porathe et al. 2018).

On a certain level one can argue that a fully autonomous ship is a simpler concept than a partly autonomous ship that requires interaction with human operators, see e.g., Endsley (2017). However, in addition to the practical and legal issues, a fully autonomous systems presents complex problems in proving that the automation is really is fully up to its task. Thus, it can be argued that it is indeed not desirable to have a fully autonomous ship.

7. Operator Control Modes

Some LOAs, particularly those based on Sheridan and Verplank’s scale of teleoperations (1978), measure *how* the operator is involved in the control of the system when responsibility for actions lays on the operator. This may range from direct control of levers to just the possibility of vetoing the automation’s proposed action. In our terminology this means that the operator is still in control and that the system is *not* autonomous. Thus, one should avoid calling this “levels of autonomy” or similar. In ISO/TS 23860 this is called operator control modes and simplified to four levels: Monitoring, strategic control, tactical control, and direct control. Monitoring may be used both in autonomous or operator controlled operation.

8. More Than One Process to Automate

In the preceding discussions it has been mentioned that there is more than one process onboard the ship that needs to be automated. Some examples are energy production, water ingress detection, stability and ballasting, fire detection and so on. Most of the focus in current literature is on sailing the ship (outlook and manoeuvring) which is arguably the most complex, as it also involves interactions with other, mostly crewed ships.

However, it is necessary to consider all processes and what level of autonomy they should have when designing new ship systems, particularly when uncrewed operation is intended. This causes some complexity in the characterization of the ship’s autonomy as compared to the individual processes’ autonomy.

One may also get into situations where different processes are controlled from different RCCs, further complicating the design.

However, the principles presented in this paper are independent of process type and should be applicable to all, although the full system aspect obviously also must be covered. That is outside the scope of this paper.

9. Conclusions

In this paper, we have argued that autonomy must mean that the operator can trust the automation to do its assigned job and, hence, that no attention from the operator is needed during autonomous operation. The control responsibility is shared between operator and automation, where only one is in control at any given time. This, however, also requires that the automation, when in control, have sufficient situational awareness to alert the operator early enough to let the operator assess the situation and safely take over the controls. Thus, there are no *degrees of autonomy*: Autonomy should be seen as a binary property, either it is there, or it is not.

However, this definition of autonomy gives rise to three or four (if full autonomy is included) cooperation modes as listed in section 5. As furthermore argued in sections 3 and 5, these modes are independent of the operators being on the ship or in an RCC. Operator control modes (section 7) can be added to show *how* the operator controls the system, when using the OE or OA cooperation modes.

Many of the existing definitions of levels of autonomy is based on similar operator control modes, sometimes modified by where in the decision pipeline the operator is placed, e.g. “in the loop” control versus “on the loop” control. Thus, one may claim that there is not anything fundamentally new in the definitions proposed in this paper. However, in our experience with safety and risk assessment of MASS, this way of defining automation capabilities, and in particular the use of the operational envelope and associated concepts, gives a clearer description of the requirements both to automation and to operators. We also believe that this way of describing interactions between operator and automation better emphasizes the necessity of proper handover of control when operational modes changes, as well as the associated timing requirements.

The proper level of trust in automation is a prerequisite for safe and efficient design and use of MASS.

Acknowledgements

The work presented in this text has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 815012 (AUTOSHIP) and No. 859992 (AEGIS).

References

CCNR. 2022. ‘International Definition of Levels of Automation in Inland Navigation’. https://www.ccnr.eu/files/documents/AutomatisationNav/DefinitionAutomatisation_en.pdf.

DNV. 2021. ‘Class Guideline - Autonomous and Remotely Operated Ships’. DNV-CG-264. DNV.

Endsley, M.R. 2017. ‘From Here to Autonomy: Lessons Learned from Human–Automation Research’. *Human Factors* 59 (1): 5–27.

Felski, A., and K. Zwolak. 2020. ‘The Ocean-Going Autonomous Ship—Challenges and Threats’. *Journal of Marine Science and Engineering* 8 (1): 41. https://doi.org/10.3390/jmse8010041.

Huang, H.M., K. Pavek, J. Albus, and E. Messina. 2005. ‘A Framework For Autonomy Levels for Unmanned Systems (ALFUS)’. In *Proceedings of the AUVSI’s Unmanned Systems North America*, 849–63.

IMO. 2018. ‘Report of the Maritime Safety Committee on Its One Hundredth Session (MSC 100/20), Annex 2 Framework for the Regulatory Scoping Exercise for the Use of Maritime Autonomous Surface Ships (MASS)’.

ISO. 2022. ‘ISO/TS 23860:2022 Ships and Marine Technology — Vocabulary Related to Autonomous Ship Systems’. https://www.iso.org/standard/77186.html.

Lloyd’s Register. 2017. ‘ShipRight Design and Construction. Additional Design Procedures. LR Code for Unmanned Marine Systems’.

Lu, Z., X. Coster, and J. de Winter. 2017. ‘How Much Time Do Drivers Need to Obtain Situation Awareness? A Laboratory-Based Study of Automated Driving’. *Applied Ergonomics* 60 (April): 293–304.

Melchers, R. E. 2001. ‘On the ALARP Approach to Risk Management’. *Reliability Engineering & System Safety* 71 (2): 201–8. https://doi.org/10.1016/S0951-8320(00)00096-X.

Ottesen, Are E. 2014. ‘Situation Awareness in Remote Operation of Autonomous Ships’. *Shore Control Center Guidelines Norway*.

Porathe, T. 2014. ‘Remote Monitoring and Control of Unmanned Vessels–the MUNIN Shore Control Centre’. In *Proceedings of the 13th International Conference on Computer Applications and*

Information Technology in the Maritime Industries (COMPIT '14), 460–67.

Porathe, T., Å. Hoem, Ø. Rødseth, K. Fjørtoft, and S. O. Johnsen. 2018. 'At Least as Safe as Manned Shipping? Autonomous Shipping, Safety and "Human Error"'. In *Safety and Reliability – Safe Societies in a Changing World*. CRC Press.

Ringbom, H.M., and R. Veal. 2017. 'Unmanned Ships and the International Regulatory Framework'. *Journal of International Maritime Law* 23 (2): 100–118.

Rødseth, Ø.J., and M. Vagia. 2020. 'A Taxonomy for Autonomy in Industrial Autonomous Mobile Robots Including Autonomous Merchant Ships'. In *IOP Conference Series: Materials Science and Engineering*, 929:012003. IOP Publishing.

Rødseth, Ø.J., L.A.L. Wennersberg, and H. Nordahl. 2022a. 'Towards Approval of Autonomous Ship Systems by Their Operational Envelope'. *Journal of Marine Science and Technology* 27 (1): 67–76.

Rødseth, Ø.J., L.A.L Wennersberg, and H. Nordahl. 2022b. 'Levels of Autonomy for Ships'. *Journal of Physics: Conference Series* 2311 (1): 012018. https://doi.org/10.1088/1742-6596/2311/1/012018.

SAE. 2021. 'Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles'. Society of Automotive Engineers.

Sheridan, T., and W. Verplank. 1978. 'Human and Computer Control of Undersea Teleoperators'. Technical Report. MIT Man-Machine Systems Laboratory, Cambridge.

Vagia, M., A.A. Transeth, and S.A. Fjerdingen. 2016. 'A Literature Review on the Levels of Automation during the Years. What Are the Different Taxonomies That Have Been Proposed?' *Applied Ergonomics* 53 (March): 190–202.

Van Hooydonk, E. 2014. 'The Law of Unmanned Merchant Shipping–an Exploration'. *The Journal of International Maritime Law* 20 (3): 403–23.

Yoshida, M., E. Shimizu, M. Sugomori, and A. Umeda. 2020. 'Regulatory Requirements on the Competence of Remote Operator in Maritime Autonomous Surface Ship: Situation Awareness, Ship Sense and Goal-Based Gap Analysis'. *Applied Sciences* 10 (23): 8751.