A BAYESIAN INFORMATION FUSION APPROACH TO NAVAL MINE-HUNTING SYSTEM OF SYSTEMS OPERATION PLANNING AND EVALUATION

Ruben Strenzke^a, Christopher Strode^a

^aNATO Centre for Maritime Research and Experimentation

Contact author: Ruben Strenzke, STO CMRE, Viale San Bartolomeo, 400, 19126 La Spezia, Italy; Tel.: +39 0187 527 447; ruben.strenzke@cmre.nato.int

Abstract: Planning and evaluating naval mine-hunting missions requires a situational picture. This picture shall include the information gathered by modern image-producing sonar sensors. Therein, a central component must fuse the information gathered by the sensors of a system of heterogeneous systems. We hence propose a novel framework for planning and evaluation for a mine-hunting system of systems. The framework calculates a geo-referenced mine presence belief by applying Bayesian statistics on detections and through-the-sensor performance self-assessment information. Thus, a mine presence probability map and an information availability map are produced; the latter showing the general confidence in either mine presence or absence. Furthermore, the framework can determine the number of future sensor looks necessary to achieve a certain target risk for vessels that are transiting the area of interest. This is done by simulation on sensor performance predictions to drive the mine presence belief to a context-derivable lower or upper threshold, which results in a sensor-specific future effort map. We describe the algorithms that generate the three before-mentioned map types from pre-processed performance assessment and target recognition data - without relying on any prior knowledge on mine density. Furthermore, we present results from simulation experiments involving systematic deviations of performance self-assessment from the actual performance. The results suggest that the geo-referenced fusion of multiple independent data collection runs improves the information availability per cell and the estimated number of mines - even if there is a constant high bias in performance self-assessment.

Keywords: Naval Mines, Mine-Hunting, Through-the-Sensor Performance Assessment, Planning, Evaluation, Bayesian Statistics, AUV, UUV

1. INTRODUCTION

Current operational approaches for the Planning and Evaluation (P&E) of naval Mine Counter Measures (MCM) operations [1, 2] were developed for manned vessels with forward looking mine-hunting sonars. These sonars do not produce lasting still images that can be analysed at any point in time later on. Each part of the area of interest (AOI) is worked on by exactly one such vessel, which will immediately pursue every detection (D) of a Mine-Like Echo with the ordered follow-up actions: classification (C), identification (ID), and neutralisation (NE).

In contrast, modern systems employ side-looking sonars that produce images for onboard or post-mission analysis. Such analyses allow Through-the-Sensor (TTS) performance assessment [3], and the result can be displayed in a 2D grid map of the seafloor [4]. Furthermore, multiple modern MCM vehicles are able to operate concurrently in the same part of the AOI - e.g. Autonomous Underwater Vehicles (AUVs) that are specialised on one or more of the mine-hunting phases (D+C, ID, NE). This means that the phases are decoupled, and detections that were established in the survey phase can be left untreated in following phases - at least for some time. As a consequence the need to fuse interim mine-hunting results arises.

In the context of our work on Planning and Evaluation for Systems (PESOS), the scenario we are interested in includes a System of Systems consisting of heterogeneous mine-hunting AUVs, which operate in the same area with different sensors (including side-looking sonars). Our goal in this work is to develop an information fusion approach that helps creating an MCM P&E situational picture. It shall be able to handle the mentioned multi-phase scenario and make use of TTS performance assessment. In addition, we show that the information produced during the evaluation process can be used as basis for decision-making and planning.

2. OVERVIEW OF THE PLANNING AND EVALUATION FRAMEWORK

The risk to vessels transiting an AOI (short: transitors) to be struck by a mine (short: risk) is the essential measure for decision making in MCM operations. Hence, the situational picture for MCM P&E must include a Mine Presence Probability Map (MP-Map), which allows deriving this risk for a specific transit route. To determine the geo-specific mine presence probability P(M) first an estimation concerning the total number of mines present in the AOI n has to be established (see Fig. 1). This leads to a probability mass function (PMF) p(n). Then P(M) can be determined from the mathematical expectation E(n) and any information collected from that cell: mine detections (passing a classification score threshold), TTS estimated probability of detecting and correctly classifying a mine (P_{DC}) , and TTS estimated false alarm rate (P_{FA}) , i.e. the probability of reporting a detection where there is no mine.

The availability of information can differ in certain parts of the AOI due to past effort, possibly contradicting results, and sensor performance. Hence, we add to the situational picture a geo-referenced information availability I and associated with it an Information Availability Map (IA-Map). Finally, we demonstrate how the planning of future MCM effort can be performed in principle. To this end, we determine the number of additional looks $L_{S,R}$ necessary to reduce the risk to a target value R with a sensor S. As input the geo-specific sensor performance prediction is needed, and the output can be displayed in a Future Effort Map (FE-Map). This feature is represented by dashed lines in Fig. 1 because it is for demonstration purpose only.

The multi-phase aspect of MCM operations is handled in the following way. All results of DC and ID looks are fused into the MP- and IA-Maps. NE task effectiveness and affordable effort has to be estimated (e.g. by a human operator). The FE-Map is generated based on this input and the target risk. If the map is generated for a survey AUV, then small spots on

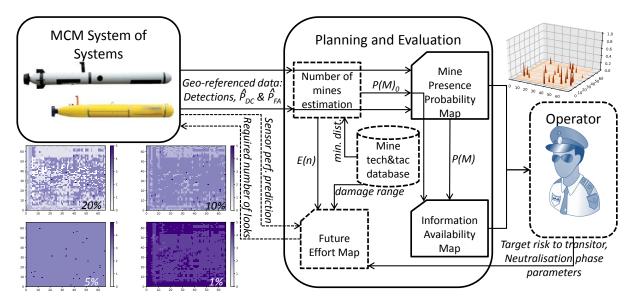


Fig. 1: Data flow in our Planning and Evaluation framework and examples of a Mine Presence probability Map (upper right) and Future Effort Maps (lower left, indicating the number of future sensor looks required to achieve a target risk, such as 20%, 10%, 5%, 1%).

the FE-Map with high number of future looks (e.g. 5 in Fig. 1) should be prioritised as re-visit locations for a different sensor with a higher predicted \hat{P}_{DC} and lower \hat{P}_{FA} (such as an ID AUV). In case NE tasks strictly depend on preceding ID tasks, ID effectiveness and affordable effort must be integrated into the NE effectiveness and affordable effort (as input for generating the FE-Map). Finally, the MP-Map plays is the indicator for NE task planning. If again ID tasks are coupled to NE tasks, then the MP-Map also has to be used for ID task planning.

3. MODEL AND ALGORITHMS

The P&E model is grid-based, and each cell can hold at most one mine and/or one detection (depending on the context). All P&E steps are completely based on Bayesian statistics.

The remainder of this chapter describes the algorithms to determine first the PMF for the number of mines p(n), secondly the cell-specific mine presence probability P(M), thirdly the cell-specific information availability I, and fourthly the cell- and sensor-specific future number of looks necessary for risk reduction $L_{S,R}$.

3.1. Estimating the number of mines present

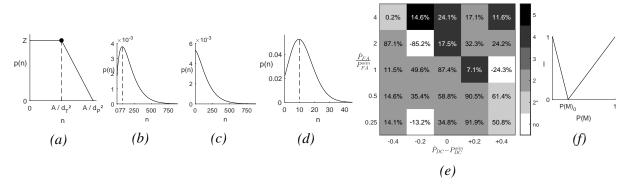


Fig. 2: (a) Initial beliefs concerning the number of mines present as PMF p(n), shaped by physical and tactical minimum mine distances dp and dT (A: area, Z: normalisation).
(b) p(n) for an AOI consisting of 50 SAS tiles, calculated after analysing only a single SAS tile containing one detection. PDC was 0.5, and PFA was 0.0005, both uniformly. E(n)=179.89.
(c) p(n) after changing PDC to 0.1 for the cell with the detection. E(n)=145.29.
(d) p(n) after analysing 50 SAS tiles, all with the parameters mentioned in (b). 16 detections in total, the most probable result if n=9. arg max p(n) = 10 and E(n)=12.36.
(e) Percentage of how close the sum of all P(M) in the AOI gets to the actual n after 5 passes over the AOI, compared to only 1 pass. 25 configurations performance estimation deviations. Also shows the number of passes that led to the first improvement over baseline (no: no improvement after 5 passes; 2*: after fusing later passes, the result was intermediately worse).
(f) Cell-specific information availability I as function of P(M) and uniform P(M)₀= E(n)/cells_{AOI}

The initial prior distribution $p(n)_{iniprior}$ holds before there is any sensor information from the AOI. It is shaped by a physical minimum mine distance d_P is used as a hard upper limit to n. In addition, a tactical minimum mine distance d_T is introduced, which is less likely to be undercut than to be overrun (linear degradation, see Fig. 2a). The resulting initial prior distribution is deemed appropriate for all kinds of MCM operations, as it does not rule out any physically possible number of mines, and it allows modelling tactical considerations.

$$p(n \mid \vec{DC}, \vec{\hat{P}}_{DC}, \vec{\hat{P}}_{FA}) \propto p(n)_{iniprior} \times P(\vec{DC} \mid n, \vec{\hat{P}}_{DC}, \vec{\hat{P}}_{FA})$$
(1)

$$P(DC_{i,j,l} \mid n, \hat{P}_{i,j,l}^{DC}, \hat{P}_{i,j,l}^{FA}) = \begin{cases} \frac{n}{cells_{AOI}} \cdot \hat{P}_{i,j,l}^{DC} + (1 - \frac{n}{cells_{AOI}}) \cdot \hat{P}_{i,j,l}^{FA} & \text{if } DC_{i,j,l} = 1\\ \frac{n}{cells_{AOI}} \cdot (1 - \hat{P}_{i,j,l}^{DC}) + (1 - \frac{n}{cells_{AOI}}) \cdot (1 - \hat{P}_{i,j,l}^{FA}) & \text{if } DC_{i,j,l} = 0 \end{cases}$$
(2)

For demonstration we created an example scenario, in which the AOI had 150×150 cells with 3×3 m size, and it contained 9 randomly placed mines. d_P was set to 3 m, and d_T to 30 m. In order to avoid any influence of the mine locations on the results, P_{DC} and P_{FA} were uniform over all cells ($P_{DC}^{sim} = 0.5$, $P_{FA}^{sim} = 0.0005$). Fig. 2b shows the updated p(n) after examining the first SAS tile containing one detection. For Fig. 2c we set a different $\hat{P}_{DC} = 0.1$ for the cell with DC = 1 to show that each the TTS context of each detection plays a role. If instead another cell (with DC = 0) was chosen to have $\hat{P}_{DC} = 0.1$, then p(n) would look similar to Fig. 2b, but E(n) = 180.23 instead of 179.89. Hence, the difference between Fig. 2b and 2c is the effect of evaluating DC in the concrete TTS context. Fig. 2d shows the outcome after analysing all 50 SAS tiles with the same attributes, but only 15 additional detections, which is the most probable outcome for an overall n = 9. Both E(n) = 12.36 and $\exp(n) = 10$ are very close to n = 9 in a solution space of $0 \le \hat{n} < 22500$.

3.2. Evaluating the mine presence probability

From p(n) it is possible to derive a uniform prior $P(M)_0$ for each cell via the expectation of n (3). After that, we can perform cell-specific Bayesian updates on P(M) for every look l ($1 \le l \le l_{max}$) that included the corresponding cell (4). The result of this iterative process can be displayed by means of an MP-Map as shown in Fig. 1 in 3D.

$$P(M)_0 = \frac{E(n)}{cells_{AOI}} \tag{3}$$

$$\forall l: P(M \mid DC)_{l} = \begin{cases} \frac{P(M)_{[l-1]} \cdot \hat{P}_{l}^{DC}}{P(M)_{[l-1]} \cdot \hat{P}_{l}^{DC} + (1 - P(M)_{[l-1]}) \cdot \hat{P}_{l}^{FA}} & \text{if } DC_{l} = 1\\ \frac{P(M)_{[l-1]} \cdot (1 - \hat{P}_{l}^{DC})}{P(M)_{[l-1]} \cdot (1 - \hat{P}_{l}^{DC}) + (1 - P(M)_{[l-1]}) \cdot (1 - \hat{P}_{l}^{FA})} & \text{if } DC_{l} = 0 \end{cases}$$

$$(4)$$

We conducted sensitivity analyses regarding the quality of the MP-Map after fusing the information from multiple passes over the AOI. To this end, we generated 25 configurations: five relations of simulated P_{DC}^{sim} to TTS-assessed \hat{P}_{DC} combined with five relations of P_{FA}^{sim} to \hat{P}_{FA} . The AOI had the same characteristics as described for the number of mines estimation experiments. For each configuration, we performed ten randomised runs of data collection/fusion. As quality measure we chose the distance reduction $\Delta \hat{n}$ [%] of the construct $\sum P(M)$ (i.e. the sum of all P(M) in the AOI) gets to the actual n. $\sum P(M)$ after five passes is compared to the baseline, which is $\sum P(M)$ after the first pass (5). These results and the lowest number of passes to improve the outcome are shown in Figure 2e. In all configurations with systematic overestimation of P_{FA} , we found that $\sum P(M)$ came closer to n than E(n) did (because p(0) became larger with every additional pass). This effect set in usually after the second pass over the AOI already, and only in one configuration after the third. After the first pass, i.e. before any cell-specific fusion, E(n) was always higher than $\sum P(M)$, but the difference was always below 0.835 [mines].

$$\Delta \hat{n}[\%] = \frac{|\sum n - P(M)_{l=1}| - |n - \sum P(M)_{l=5}|}{|n - \sum P(M)_{l=1}|} \cdot 100$$
 (5)

3.3. Evaluating the information availability

The goal of an MCM operation is to localise mines, which means to maximise the certainty concerning mine existence or absence at any position. Therefore, the information that we want to gather is directly related to P(M). We base the geo-referenced information availability measure on the difference of P(M) and $P(M)_0$. Hence, for cells never looked at I=0, because their P(M) was never updated by (4) to differ from $P(M)_0$. For obvious reasons, I reaches the maximum 1 in case of total certainty regarding mine presence or absence, i.e. $P(M)=1 \lor P(M)=0$. We apply two linear scales for $P(M)>P(M)_0$ and $P(M)<P(M)_0$ (see Fig. 2f).

We conducted sensitivity analyses regarding the effect on information availability by fusing the information delivered from two passes over the AOI. To this end, we used the same parameters and the 25 configurations for the deviations of TTS assessment as in Section 3.2 and performed ten randomised runs of data collection per configuration. The fusion was shown to improve I averaged over all cells by either 0.09, 0.21, or 0.25 - depending solely on \hat{P}_{DC} .

Table 1 shows the possible developments of P(M) and I over two looks, assuming $P(M)_0$ remains unchanged. Again, the same parameters were used. When P_{DC} =0.5 \land $P_{DC} \gg P_{FA} \land P_{DC} > P(M)_0$, then every DC=0 event lets P(M) fall to \approx 50% of the previous value, hence

Table 1: Results for P(M) and information availability I after one and two looks on a cell with different detection and classification outcomes \vec{DC} . \hat{P}_{DC} =0.5, \hat{P}_{FA} =0.0005.

	$\vec{DC} = (0)$		$\vec{DC} = (1)$		$\vec{DC} = (0,0)$		$\vec{DC} = (0,1)$		$\vec{DC} = (1,1)$	
$P(M)_0$	P(M)	I	P(M)	I	P(M)	I	P(M)	Ī	P(M)	Ī
0.0004	0.0002	0.4997	0.2858	0.2859	0.0001	0.7497	0.1668	0.1669	0.9975	0.9979

decreasing the distance of I to 1 by $\approx 50\%$. In contrast to that, cells with true or false alarms end up with $P(M) \gg P(M)_0$, so that $I \approx P(M)$. For DC=0 events, I develops over multiple looks like $1-\underline{P}_{DC}$ (i.e. P_{DC} aggregated over multiple looks) as long as $P_{DC}\gg P_{FA}$ and $P_{DC}>P(M)_0$. Although I could therefore in principle be used as the maximisation objective for combined coverage and detection revisit planning, we propose to plan on a target risk in the following.

3.4. Estimating the future survey and identification effort

If it is possible to estimate future sensor performance, e.g. from past performance, then we can determine for every cell the future number of looks $L_{S,R}$ necessary to probably achieve a certain target risk with this sensor. Risk reduction is two-fold: For empty cells we require a confidence on mine absence; and for mines we want to reduce the risk of remaining unneutralised (i.e. it either does not become the target of any NE attempt or all attempts fail).

We define $thresh_{empty}$ as the value to be undercut by the P(M) that is predicted after simulating future effort. Its value can be determined as shown by (6) to (7): via the target risk R_{total}^{empty} and the number of cells in which the presence of a mine would pose a threat to the transitor (cells to cross length-wise c, potentially dangerous cells width-wise per crossing step m, determined by mine damage range $r_{damage[cells]}$ to both sides of the ship, hence \times 2).

$$m=2 \cdot r_{damage[cells]}$$
 ; $R_{step}^{empty} = 1 - (1 - R_{cell}^{empty})^m$; $R_{total}^{empty} = 1 - (1 - R_{step}^{empty})^c$ (6)

$$m=2 \cdot r_{damage[cells]} ; \quad R_{step}^{empty} = 1 - (1 - R_{cell}^{empty})^m ; \quad R_{total}^{empty} = 1 - (1 - R_{step}^{empty})^c$$
(6)
$$thresh_{empty} = R_{cell}^{empty} = 1 - \sqrt[c-m]{1 - R_{total}^{empty}}$$
(7)

When a cell's P(M) surpasses thresh_{mine} (which is higher than thresh_{empty}), this serves as an indicator to perform NE attempts. Because the number of NE attempts available is usually limited by certain factors, a high thresh_{mine} is needed to achieve a low R_{total}^{empty} . This can be explained by filtering out false alarms that would otherwise be targeted by NE tasks, thereby reducing the number of attempts on actual mines and increasing remaining risk. As first step to determine $thresh_{mine}$, in this uniform approach we demand the risk posed by cells with mines to be reduced to a similar value as that of empty cells, and the remaining risks depends on the probability of not trying or (repeated) failing of NE attempts - hence (8), where $f_N \in \mathbb{R}_{>0}$ is the average number of attempts per cell where we suspect a mine, i.e. where $P(M) > thresh_{mine}$. (9) describes how to determine f_N in case either less than one attempt is needed on average or more than one. In realistic examples of sparsely populated minefields, R_{cell}^{mine} will already be very low compared to the product of reacquisition and NE success probabilities $P_{RN} \cdot P_N$. Then f_N is high, and hence it is better to target for successful NE only of mines along a single route, which can be either generic or specific. This is described in (10), where $cells_{alongRoute}$ refers to the specific channel made up geometrically by $c \times m$. As shown in (11), thresh_{mine} can then be determined by distributing \hat{n} as P(M) over the number of cells that will be visited for NE v_N , which can be derived from f_N and the number of NE tasks available t_N .

$$R_{cell}^{mine} \equiv R_{cell}^{empty}$$
; $R_{cell}^{mine} = 1 - f_N \cdot P_{RN} P_N$ if $f_N \le 1$; $R_{cell}^{mine} = (1 - P_{RN} P_N)^{f_N}$ if $f_N > 1$ (8)

$$\implies f_N = \begin{cases} \frac{1 - R_{cell}^{empty}}{P_{RN}P_N} & \text{if } R_{cell}^{empty} \ge 1 - P_{RN}P_N \\ \frac{\log(R_{cell}^{empty})}{\log(1 - P_{RN}P_N)} & \text{if } R_{cell}^{empty} < 1 - P_{RN}P_N \end{cases}$$

$$(9)$$

$$\hat{n}_{alongRoute}^{generic} = \frac{E(n) \cdot c \cdot m}{cells_{AOI}} \quad ; \quad \hat{n}_{alongRoute}^{specific} = \sum_{\forall cells_{alongRoute}} P(M)$$

$$thresh_{mine} = \frac{\hat{n}_{alongRoute}}{v_N} \quad ; \quad v_N = \frac{t_N}{f_N}$$

$$(10)$$

$$thresh_{mine} = \frac{\hat{n}_{alongRoute}}{v_N} \qquad ; \qquad v_N = \frac{t_N}{f_N} \tag{11}$$

We now determine the number of future sensor looks $L_{S,R}$ that with at least a probability of p_{min} (e.g. 50%) achieve that: P(M) undercuts thresh_{empty} if the cell is empty, and P(M)surpasses thresh_{mine} if there is a mine. For every cell with thresh_{empty} < P(M) < thresh_{mine}, $L_{S,R}$ is iteratively incremented from 1 upwards with abort criteria as described below. In every iteration two Bernoulli experiments are made with predicted P_{DC} and P_{FA} for a specific sensor in the geographical context: It is simulated that the cell is empty (with the success case DC=0), and it simulated that there is a mine (with the success case DC=1). In each simulation all combinations of numbers of DC=0 and DC=1 events are processed by iterative Bayesian updating (as in (4)), wherein the current P(M) of the cell serves as prior. We then sum up the occurrence probabilities of all combinations by which the predicted P(M) passes the corresponding thresh. If that sum surpasses p_{min} for the mine and empty case, then $L_{S,R}$ has been found for that cell. Otherwise the next higher $L_{S,R}$ has to be tested.

Thus, for a specific sensor and a certain target risk value, an FE-Map as displayed in Fig. 1 can be generated. It shows in where how much effort is needed to reduce risk to a target value, but it is neither suited to replace track planning, nor is it suited as input for track planning. This is because a certain track pattern leads to different numbers of looks for different cells (due to overlaps and gaps between SAS tiles) and at the same time to range-dependent effects on P_{DC} per cell [4]. However, inside this framework, a track planner could iterate over its parameters, thereby influencing $L_{S,R}$ and P_{DC} to reach the thresholds.

CONCLUSION, RELATED AND FUTURE WORK

We have described a framework that performs the whole MCM P&E loop, with the exception of track/path plan generation. Its novelty is that it works with detailed geo-referenced maps - in addition to the classical statistical measures. By means of the experiments presented, we have shown that even if the AUVs' performance assessment has a high deviation from actual performance, data fusion from multiple passes or vehicles has a positive effect on information on mine presence or absence. With the FE-Map, we demonstrated how sensor performance prediction can be used for the planning of MCM effort based on risk to transitors.

The TTS performance assessment required by our approach has been performed on the estimated probability of detection and correct classification of a mine already [4]. In addition, similar to [6] we have proposed that TTS performance assessment be performed also for the false alarm rate. We applied a grid-based approach, as it is done in [7]. Their concept of a "search channel" treats detections and non-detections for every grid cell over time as Bayesian update on the associated mine presence belief. Although our approach is similar in that, it differs through introducing: first, a solution for the important question of the prior mine presence belief, and secondly, mine presence probability thresholds for planning on risk.

Future work includes analysing the impact of navigational error (i.e. probability distributions of the AUV position) on the geo-referenced mine presence probability. There is a cell-size effect when treating navigational error and when calculating the risk - both needs to be taken into account. The proposed approach can be improved in terms of computational speed by finding a non-optimal unbiased estimator. The estimation of false alarm rates and the possibility of reproducible false alarms in the survey phase (e.g. rocks with mine-like shapes) should be target of future research in the context of real sensor data. Related to that is the analysis of false alarm geographic distribution patterns, especially in comparison with realistic mine-laying patterns. Also, the exploitation of classification scores should be considered in the likelihood function - in contrast to treating every detection that passes a classification threshold in a binary way. In that context it is also relevant that the scores might systematically depend on aspect angle. Furthermore, the FE-Map points to the requirement that a future planning component needs to perform mixed-sensor planning. If the FE-Map is generated for an ID AUV with P_{DC} high and P_{FA} low enough so that all cells show the future number of looks is 1, then ID task optimisation is pointless. Instead, many cells might need to be part of an additional survey before planning ID tasks, if the survey sensor has a wider field of view. Related to that is also the unsolved topic of sensor performance prediction models (i.e. dealing with past performance, environmental conditions, aspect angle, range etc.). Finally, to apply the framework to different operating navies, standards for data exchange need to be defined and tested under real world conditions.

5. ACKNOWLEDGEMENTS

This work was supported by NATO ACT. The authors have taken up ideas exchanged and approaches discussed in the context of the MCM P&E Joint Research Project, especially from FFI (Norwegian Defence Research Establishment) and the German Navy.

REFERENCES

- [1] Washburn, Alan R. "Mine Warfare Models." (2007).
- [2] Cramer, Megan. *Understanding information uncertainty within the context of a net-centric data model: A mine warfare example*. Naval Sea Systems Command Washington DC Mine Warfare Program Office (PMS 495), 2008.
- [3] Harris, Mike, et al. "AQS-20 Through-the-Sensor (TTS) performance assessment." *OCEANS 2005. Proceedings of MTS/IEEE*. IEEE, 2005.
- [4] Gips, Bart, Strode, Christopher, and Dugelay, Samantha. "Residual Risk Maps for Performance Assessment of Autonomous Mine Countermeasures using Synthetic Aperture Sonar". *Proceedings of the Insitute of Acoustics*. 2018.
- [5] Lynch, Scott M. *Introduction to applied Bayesian statistics and estimation for social scientists*. Springer Science & Business Media, 2007.
- [6] Geilhufe, M., Connors, W. A., and Midtgaard, Ø. "Through-the-sensor performance evaluation for modern mine hunting operations." *Proc. Undersea Defence Technology*. Lillestrøm, Norway. 2016.
- [7] Baylog, John G., and Wettergren, Thomas A. "A ROC-Based Approach for Developing Optimal Strategies in UUV Search Planning." *IEEE Journal of Oceanic Eng.* 2017.