Evaluating the Data Visualization for Demanding Marine Operations

Xu Zhuge, Hao Wang*, Girts Strazdins
Big Data Lab, Faculty of Engineering and Natural Sciences,
Norwegian University of Science and Technology, Aalesund, Norway

Abstract— **Complexity and specificity of monitoring sensor data from marine operations brought a majority of challenges of how to evaluate the visualizations of such sensor data. It's urgently needed to develop and implement new evaluation methods to assess the visualization of maritime data. In this paper, we explore both quantitative and qualitative evaluation approaches to define evaluative metrics of our interactive visualization techniques, and present a case study carrying out by domain researches in real setting. This work is an important part to complement the visual analytics for maritime operations.**

Index Terms—**Maritime operations; evaluation; visual analytics; data visualization.**

I. Introduction

New sensor technologies are being adopted to monitor demanding maritime operations. The monitoring data collected can also be used to assist ship designers in analyzing and optimizing vessels. However, the large volume of the monitoring data brought many challenges. One significant challenge is the increased complexity derived from various technologies and integrated operations, causing *Information Overload Problem* (IOP) [1]. There is no effective way to transfer from data to knowledge, to support decision making. Another significant challenge is the poor quality of the raw sensor data. The data is multifaceted and spatiotemporal, which generally has a lot of errors, gaps and overlaps.

In order to address such challenges, in the Big Data Lab at NTNU Aalesund (BDL), we have proposed a *Visual Analytics* (VA) framework for maritime operations [2]. The conceptual framework is depicted in Figure 1. Data integration combines the data from diverse sources and provides a unified view of the data. Data cleaning aims to improve data quality, e.g., define and clean the noises, detect and eliminate the gaps and overlaps among data. The visualization is implemented interactively based on the processed datasets and the hypotheses generated from domain knowledge.

Guided by the VA framework, papers [2][3][4] presented our initial results in data integration and visualization, efficient pattern identification, and prediction respectively. We have obtained real monitoring data from the HEalth MOnitoring System (HEMOS) by Rolls-Royce Marine AS, covering three years' operation by one vessel including high frequency machinery and low frequency vessel behavior data. We have built an initial proof-of-concept prototype [2] to visualize the behavior of the vessel, and allow ship designers, crewmembers and domain experts to easily explore the correlation coefficients, the temporal and spatial aspects of the data. In this paper, we focus in the assessment and evaluation of our data visualization, which is a highly challenging task itself [5].

Data visualization becomes more established as an important topic for data management in different domains. As a result, it becomes increasingly important that the research outcome in such field can be validated. It has been suggested that current evaluation methods are not convincing enough to prove the usability of the visualization, and to encourage widespread adoption of the visualization tools. Reasons given generally include unsatisfied samples, incorrect evaluative methods, misguided, or simple evaluative tasks.

For the data visualization of maritime operations, it is particularly difficult to establish the evaluation methods and define the evaluation tasks because of the complexity of the data collected. One major challenge in designing the evaluation methods is that the chosen method should be able to evaluate whether the visualization tools are accessible to diverse users regardless of their backgrounds and technical disadvantages [6]. For this reason, we use both quantitative and qualitative evaluation approaches [7]. These two approaches are inspired by the empirical research from the area of human-computer interface (HCI). When designing evaluation methods in the two approaches, we have addressed the issues such as how to obtain an appropriate sample of participants, how to select suitable evaluation methodologies, and how to analyze the outcome.

Another major challenge in designing the evaluation methods is the definition of the evaluation metrics that assess the information value and human interaction [8]. The value of information is one of the most important criteria of visualization [9]. Thus, the designed metrics should be able to determine the value of information that extracted from the visualization and to verify how the extracted information and the user interactions can support decision-making.

We have designed both quantitative and qualitative evaluation methods and performed a case study in which participants carrying out specific data analytic tasks. We have designed a fair amount of evaluative tasks; we have organized an in-situ interview for group of domain experts; we analyzed and summarized the feedback from interviewees. For the quantitative method, we designed the experimental setup and

* Corresponding author, hawa@ntnu.no, Norwegian University of Science & Technology, Postboks 1517, 6025 Aalesund, Norway

the standards for participants. A questionnaire including representative tasks which measured the participants' correct rate and response time was devised. For the qualitative method, user feedbacks were collected with open-ended questions and discussions during the group interview.

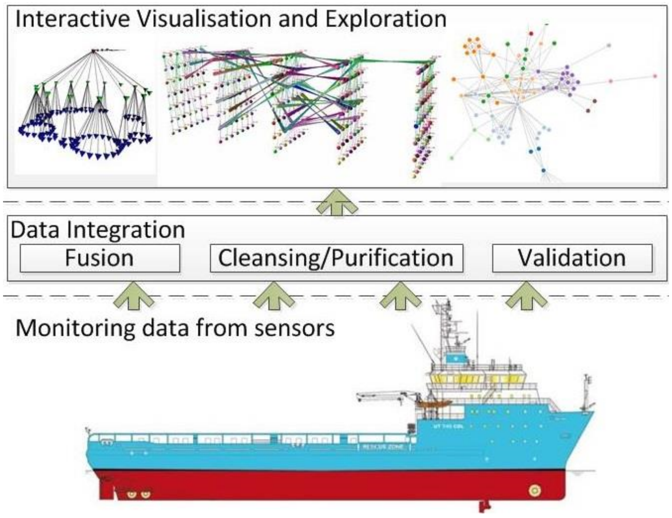

Figure 1. Integration and visualization of monitoring data for demanding marine operations

The remainder of this paper is structured as follows: Section II reviews existing literature on evaluations of data/ information visualization. Section III presents two different approaches to evaluate the visualization. Section IV describes our case study, with details in the process, the environment and the assessment criteria. In Section V we discuss the evaluation results of the case study. And Section VI concludes the paper and indicates some future directions.

II. BACKGROUND

The evaluation of users' reasoning capabilities and sense-making process is an indispensable phase in the practical utilization of visual analytics in new application domains and research areas, e.g., marine research. Scholtz [10], Thomas and Cook [11], Thomas and Kielman [12] have indicated the importance of new methodologies and metrics to support the researchers measure the usability of visualization and understand the impact of visual analytics process.

Plaisant [5] has summarized the current evaluation practices and identified three main problems that designers met during evaluating the data visualizations:

1. Practicality: Evaluative methods should match the visualization tools with users, tasks and real problems. The utility of visualization needs to be demonstrated in a real setting, e.g., an operable application platform, an adoptable test environment, and a group of participants.
2. Improvement: Tasks that located and identified from experiments should not be simple. Specific aspects of empirical studies could be improved, e.g., the evaluative tasks enable users to look at the same data from different perspectives.
3. Universality [6]: Users from diverse backgrounds must understand the visualization tools according to the guideline of evaluative tasks. For instance, the designer should considerate that the tasks are accessible to general participants from other background.

Other problems include: 1) the difficulty of obtaining an appropriate sample of participants, e.g., it's hard to obtain domain experts' time; 2) the user may make an existing interactive visualization technique as a benchmark against the prototype in an involuntary way if he/she is familiar with one existing software, and this behavior may affect the evaluative feedbacks.

In order to solve such problems, Carpendale [7] presented two different types of evaluation methods: *quantitative* approach and *qualitative* approach. The goal of Quantitative evaluation is to define the representative tasks [13] that participants carry out, and standardize the evaluation metrics, which are suggested as correctness of response, time to complete representative tasks, and user satisfaction regarding the visualizations provided [10]. On the other hand, quantitative assessment is complemented with a qualitative evaluation that carried out with domain researchers and experts. Qualitative methods toward achieving a richer understanding by using extended approaches. These approaches require considering the internal and deep-seated factors that influence the visualization, and the performance and development of it [14]. Carpendale [7] mentioned that qualitative evaluation typically involves two approaches: interview [15] and observation [16]. Both techniques could be used in visualization evaluations, and be used to collect feedbacks from domain experts.

One key part of visualization assessment is sense-making, which provides a theoretical framework for understanding the analytical reasoning process, maximizing human perceptual and reasoning capabilities in complex and dynamic situations. Furthermore, Duffy [17] has defined the sense-making as "how people make sense out of their experience in the world". Russell et al. [18] have introduced the sense-making loop which involves the iterative process of hypotheses validation and evaluation of visualizations. Therefore, in evaluation process, sense-making is an important phase that helps users acquire insights, and make decisions based on their knowledge or experience.

Although data visualization is becoming more established, technologies of validation are still immature, especially in the marine community. Only few researches were found in public w.r.t. evaluating the visualizations of marine geography [19] [20]. As a result, the research in this paper (evaluating the data visualization for marine operations) represents a new area in maritime community.

III. ANALYTIC EVALUATION

Data collected from HEMOS presents a set of features [2], w.r.t. multivariate, multimodal and spatiotemporal property. In order to define the metrics to assess such features and support decision-making, the evaluation must involve two tasks:

1. Understand normal behaviors of marine operations, e.g., GPS information of a vessel during one year.
2. Validate the hypotheses generations based on users' domain knowledge, e.g., the correlations between several modules in a vessel.

A. Quantitative Approach

Providing insight is considered as the main purpose of data visualization and visual analytics. A significant measure criterion is the quantity of insights generated from the visualization and the value involved in the insights. As an effective way to quantitate the characterization factors of larger samples, the quantitative evaluation method is established to illustrate how the study will be conducted and what tasks the participants will carry out.

Quantitate evaluation involve a strict process of hypotheses generating and development, identification of independent variates, definition and measurement of the correlations between dependent variates, and the statistics methods to quantitate the metrics which make the results to be trustworthy. The results should provide sufficient evidence to certify the effectiveness, usefulness and performance of visualization. Furthermore, measurement of user interactions is also considered as one part of the quantitative evaluation of visual analytics. The representative tasks should enable researchers collect useful information about user reactions of their interaction with the visualization.

In order to solve the challenges in quantitate evaluation, Koua et al. [8] have formulated the exploration tasks and proposed a list of visual operations as a general solution to assess the usability of visualizations. However, according to the specificity and complexity of marine monitoring data, the goals and tasks of assessment are different. Hence the hypotheses generated from domain knowledge are different. Based on Koua's research, we implemented several modifications and provided ten visualization operations, which are defined as follows:

1. Identify: to establish a set of characteristics that can distinctly represent an variate;
2. Associate: to link or join in a relationship between several variates;
3. Correlate: to establish a direct connection and provide the correlation coefficient between two variates.
4. Categorize: to place one specific group/module into a classification (e.g., classification can be defined by one color, shape, position, etc.);
5. Distinguish: to recognize the differences or distinctions;
6. Cluster: to join into groups of the same, similar or related variates;
7. Distribution: to describe the modules in an overall way, which presents how modules distributed and how one module effects the entire system;
8. Rank: to sort the variates and give an position of specific variate with respect to other variates;
9. Range: to define the range of value and determine the position in the certain range;
10. Compare: to compare behavior among different modules in spatiotemporal condition.

To effectively recognize the vessel behavior and the correlation between variates, visualization operations need to be carried out with background information. As a result, the operational visualization tasks are divided to several specific tasks or "low-level" queries, which are based on maritime research knowledge and explored in the study.

Table 1 presents the visual operations, which contain a set of real operational visualization tasks. Each operational visualization task involves one or two specific user tasks that are presented to illustrate the detail tasks, and are to be implemented by the participants in case study. We made a list of correct answers and provide a standard of correctness from grade "0" to "100%". Response time is also considered as an important metric. Time to complete each task is to be recorded in the case study, and to be analyzed as quantitative metrics for measuring the speed that the users generate insights and the extent that how easily the users can work with the interface.

Table 1. List of visual operations, operational visualization tasks and specific tasks for evaluation.

Visual operators	Operational visualization task	Specific task explored in the study
Identify	Identify one variate	Identify "Latitude" at one time point
Associate	Form relationships between variates	Form relationship between "Rotation Speed" and "Torque: Main Bridge"
Correlate	Discern variates that have high correlations	Discern 3 couples of variates that have high correlation coefficient ($\geqslant$ 0.8)
Categorize	Define models on the display, indicate the physic meaning	Categorize a model in the prototype and tell the physic meaning
Distinguish	Distinguish how a target variate effects other variates and how it significantly impacts one module	Distinguish how does variate "Temperature" effects other variates, and distinguish which module it significantly impacts
Cluster	Find groups of correlated variates	Find 3 groups of correlated variates
Distribution	Describe the overall pattern	Summary that how "Temperature" module impacts the vessel system
Rank	Indicate the best and worst cases of one variate	Find the moment that vessel has the highest and lowest velocity
Range	Indicate the value of variate in a certain range	Indicate the value range of variate "speed" at one time point
Compare	Compare values at different time point	Compare variate "speed" at two different time points

B. Qualitative Approach

In order to complement the quantitative assessment, qualitative approach is commonly used to evaluate the usability and performance of visualization tools. The main task of the qualitative evaluation was used as a synergistic approach to collect domain experts' impressions, insights and suggestions based on their expertise.

Qualitative evaluation has been proven an effective approach to achieve richer understanding of the usability and development of visualization [21]. Studies in qualitative evaluation normally occur in real settings, they improve the understanding of existing practices and may inspire new ideas for the prototype.

To evaluate our data visualization, we designed the qualitative assessment, which consists of: (1) one case study including questionnaires, on-site operations and response collection; (2) an in-situ interview with a group of persons with marine research background; (3) evaluative criteria to summarize the feedback from participants. One major objective of carrying out the experiments with those maritime researchers is to obtain their impressions and opinions of the visualization, and confirm that whether the visualization could be used to complete the specific tasks presented in Table 1. The group interview aims to confirm the opinions and impressions of maritime researchers and complement the quantitative results. Four objectives are established for the group interview: (1) provide suggestions of user interactions; (2) discuss the impressions generated from their background knowledge; (3) obtain potential information of the vessel; and (4) predict the future of maritime data visualization. The procedures of qualitative evaluation, the responses from participants and the conclusions will be discussed in the subsequent sections.

IV. CASE STUDY

Since there are few previous studies regarding how researchers visualize the monitored data from marine operations and how visual analytics can support the visualization, we carefully devise a case study to assess our visualization prototype. The objective of the case study is to provide a description and analysis of how the domain experts study the ship normal and abnormal behavior, supply and assess the hypotheses of variate correlations, provide suggestions for future works, and improve the usability of data visualization.

In order to achieve the quantitative evaluation goals, we designed a user study to implement the case study and establish how the participants carry out the specific tasks discussed in Table 1. The user study includes (1) a questionnaire with 17 specific representative tasks, and (2) an in-situ interview organized on the campus of the Norwegian University of Science and Technology (NTNU).

A. Participants

There were 13 participants in the case study, including 8 master students from the Maritime Technology and Operations department of NTNU; 2 master students from International Business and Marketing department of NTNU; 2 exchange master students majoring in Computer Science from China; and one PhD student from Computer Science department of University of Oslo. Among them, 2 participants are female. The average age of all participants was 27.4 years. All participants have normal or corrected to normal vision. 13 participants were divided to 2 groups, one was reference group that consists of 8 participants from maritime department of NTNU, and the other consists of 5 participants from different majors.

Ideally, a larger number of participants and groups with equal participants could provide more fairness, trustworthiness and comprehensibility. However, with the reality constraints, we managed to carry out the user study and obtain useful qualitative and quantitative results and many helpful feedbacks.

B. Study Setup

The test session is on individual basis, and each participant is to work individually in a room, with no disturbance and minimum noise. The room is equipped with one laptop, mouse, two displayers, and one recording monitor, depicted in Figure 2. One questionnaire and one pencil with eraser are prepared before each participant enters the test room.

The participant is to use the visualization prototype to carry out representative tasks and analyze the real vessel behaviors. The paper questionnaire specified the tasks to be carried out, it includes 17 tasks, 4 discussions and suggestions for the prototype. The tasks are varied from specific questions to the questions that have more open insights. A group interview regarding user reactions, questions about satisfaction and suggestions for the improvement was organized after the individual questionnaire.

C. The questionnaire

The evaluation approach assesses the visualization prototype in human-computer interaction and vessel behavior representations, in order to help the researchers understanding the usability and development of maritime operations visualization. The tasks are divided to small questions from different perspectives, for example, "How does temperature variate affect other variates? Give the explanation based on domain knowledge." (about the understanding of correlations between vessel modules), and "Describe the trajectory of the vessel. And tell when and where did the vessel port in harbor?" (about the understanding of vessel behavior).

Analytics of the representation with/without maritime knowledge is provided to the participants as an aid in performing the tasks. Correctness of response is the first criterion of the extent of information extraction. Main goal of the first criteria is to measure whether the visualizations enable user understand the vessel behaviors and correlations between variates.

The second criterion is the time spent on completing each task. Therefore, both groups are compared regarding the correctness of their responses and the time to complete tasks. Time is recorded to observe the decision-making of each participant and the extent of operating the user interface.

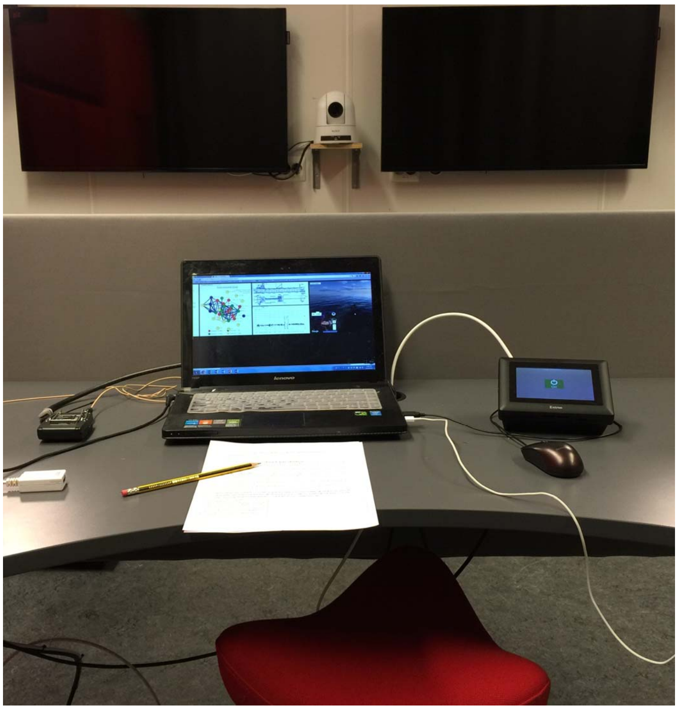

Figure 2. User study setup for participants carrying out evaluation tasks

The third criterion, user reactions or feedbacks, assess the appropriateness of the visualizations and the extent to which representations meet the goals and expectations of participants. In addition, user reactions are collected with open-ended and rating questions, thus they can interpret the level how the participants understand the data visualization. An in-situ interview is held to collect the user reactions for each participant. Comments and suggestions based on domain knowledge are collected to be one part of user feedback.

V. RESULTS AND DISCUSSTION

In this section we analyze and discuss the results of case study. We used the three above-mentioned criteria to assess the visualization prototype: correctness of response, time to complete each task and user reactions. As a part of quantitative evaluation, statistics methods are used to quantitate the criteria.

A. Correctness of Response

Standard for assessment of participant's answer was created, the correctness rate is defined from 0 to 100 percent, $C \in [0,100\%]$. Equal means of the results is analyzed by *two sample t-test* (significance level $\alpha = 0.05$, null hypothesis implies that the samples come from populations with equal means, $p = 0$ represents the two samples are definitely diverse while $p = 1$ represents that the two samples are equal). Table 2 depicts the average correct rate and response time for both groups of participants to complete evaluation tasks. The average data would assist researchers in analyzing the process of decision-making and the extent that users following the visual operations and understanding the information.

Figure 3 depicts the result of the analysis of correctness metric. The blue bars represent the results of the group with maritime background. The yellow bars represent the results of baseline group, without maritime background. Questions 1 to 7 (left side) are questions with specific answer from general domain. Questions 8 to 17 (right side) are open-ended and rating questions related to maritime domain.

The analysis of correctness of response shows that the participants from both groups have followed the visual operations and have obtained basic information from the visualization prototype (100% correct rate of question 1,5,6,7,9 and 10). The t-test analysis of the correctness of general questions (1 to 7) reveals that there is no significant difference between two groups ($p = 0.9333 > \alpha$ for question 2; $p = 0.6783 > \alpha$ for question 3; $p = 0.787 > \alpha$ for question 4). However, the more open insight the questions are, the more decrease is seen in the correctness of the second group. Even the participants without maritime background did not provide any correct answer in question 16.

The participants from both two groups did well in general questions ($C_1 = 87\%$ and $C_2 = 86\%$). Both groups could extract basic information from the visualization, e.g., variable definition, time information, geographic information, etc. The results reveal that the visualization prototype helped general researchers extract basic maritime information from the sensor data, thus support users understand normal marine operations.

However, when the questions became more open insight-like, the participants from the non-maritime group could not achieve deeper understanding and make effective decisions ($C_2 = 59\%$). Hence, they had no capabilities to accomplish sense-making tasks without maritime knowledge. Meanwhile, the average correctness rate of group 1 was still at a high level ($C_1 = 85\%$) The visualizations were helpful for maritime researchers to effectively and efficiently generate more insights and extract valuable information from the monitored data.

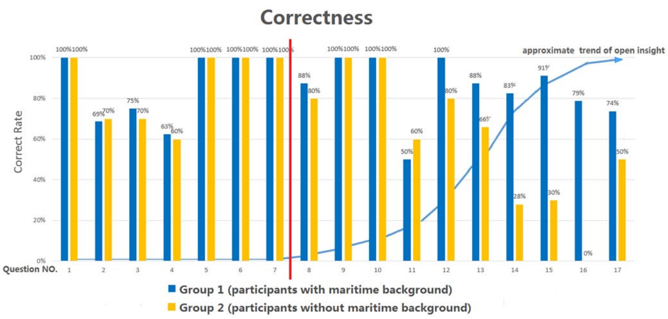

Figure 3. Average correctness of response per question.

B. Time to Complete Tasks

Second criterion, time to complete tasks, is used to support the evaluation measuring the visualization performance. In order to observe the decision-making process of participants, the analysis of results is divided into two portions: time to complete general questions, and time to complete open-ended questions.

Figure 4 and Figure 5 depict the average time for two groups to complete the questions contain general tasks and to complete the questions that have more open insights, respectively. In Figure 4, the t-test analysis of the average time indicates that no significant difference between two groups was found for the general questions. The major difference question 1 to 7 is $p = 0.124 > \alpha$ (for question 5). Time spent by participants from both groups to complete general questions is close, average time spent by group 1 is 77.6s while for group 2 is 77.9s. The results reveal that general researches spent similar time as domain researchers to complete representative tasks that only required visual operations. Human interactions enabled users effectively and efficiently carry out all the evaluation tasks.

However, there is a significant difference when the questions have more open insights. The t-test analysis of the average time depicted in Figure 5 reveals the significant difference, $p=0.00335 < \alpha$. Participants from the first group spent more time completing the tasks (103.6s). The screen recorder and audio logging indicate that the participants with maritime background spent much time considering and explaining the insights generated from their domain knowledge, e.g., "temperature" is one of the most important variables because energy loss causes temperature difference. Instead, average time spent by the second group was 58.9s, which was approximately half of the time spent by the first group. Participants with domain knowledge needed more time to finish the tasks; they performed better and gave more detailed reasons for variable correlations and vessel behaviors; and they were more confident in their decisions. On the other hand, for participants without domain knowledge, they focused on how to complete visual operations and how to verify the performance of user interactions instead of explaining professional opinions.

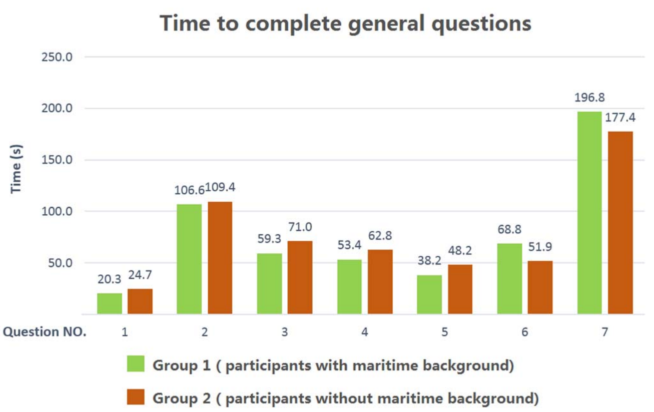

Figure 4. Average time for participates to complete first 7 questions.

C. User Reactions

In order to achieve the objectives of qualitative assessment, an in-situ interview for each participant was organized, and user reactions generated from such interviews were analyzed to assess the usability of the visualization.

Regard to the capabilities of visualizations that supporting maritime researches, interviewees from maritime background found that the visualizations were useful and which had developmental potentiality. The interviewees from ship design department mentioned that "detection, definition and visualization of the correlations between variates of the vessel are helpful to the ship design". Two interviewees found that the visualization could help with the ocean environments and risk-averse operations ("for 3D animation, the weather data could be involved to support the observation of the vessel behaviors").

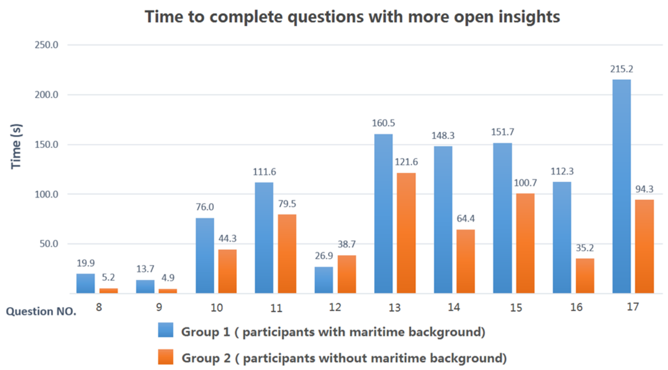

Figure 5. Average time for participants to complete last 10 questions.

Table 2. Summary (average) of correct rate and time to complete

Average Data Analytics	Correct rate of Question NO. 1 to 7	Correct rate of Question NO. 8 to 17	Time to complete NO. 1 to 7	Time to complete NO.8 to 17
Group 1	87%	85%	77.6s	103.6s
Group 2	86%	59%	77.9s	58.9s

In addition, interviewees from different domains found that the visualizations could help them understand normal vessel behaviors and basic maritime knowledge. They also provided suggestions, e.g., one interviewee from statistics background suggested that "the use of statistics to obtain correlations is helpful, but it should be updated, e.g. generating more clear model".

The feedback from the representations and user interactions were positive, since they support the understanding of the vessel system and vessel behaviors, and increase the trust of the outcomes. On the other hand, the interviewees argued that some optimizations of the visual representations might be implemented. For example, some interviewees suggested that "to create an interactive timeline that could observe the variates changing with real time", and several interviewees commented that "the parameters were hard to find in the chord diagram and line chart, the visual representations should obtain more human interactions to support users completing operational visualization tasks."

The final issue discussed was the feasible future works and the predictable possibilities of the data visualization of maritime operations. The interviewees discussed the issue of updating and modifying such visualizations and which parts should carry out such tasks in the future. The interviewees encouraged the use of visual representations to "reduce reaction and decision time". They also mentioned the possible future works that using visualization to find the internal relationships between vessel modules, guaranteeing maritime safety, and reducing the operating costs.

VI. CONCLUSION AND FUTURE WORK

In this paper, we introduce a guideline of evaluating the visualization of monitoring sensor data from maritime vessel operations. Following the principles in our discussions, we present a case study for implementing the quantitative and qualitative assessment methods.

As part of the ongoing research project, we aim to explore further to improve the evaluation. Comparison with other existing visualization tools (benchmark) is a viable approach to assess the usability of the prototype. Several evaluation metrics such as 2D model representation, 3D model representation, and user interactions could be set up to support the comparison. Furthermore, the number of participants should be increased as a sufficient sample size entails a better quantitative assessment. In addition, in-depth interviews [22] with authoritative experts (e.g. professors from maritime research institutions) should be considered to support the qualitative assessment.

ACKNOWLEDGEMENTS

The authors would like to thank for the valuable inputs from other members from BDL and Houxiang Zhang and Guoyuan Li from the Faculty of Maritime Technology and Operations of NTNU Aalesund; Rune Garen, Leif R. Solaas, Thomas Oksavik, Krzysztof Swider, and Are Folkestad from Rolls-Royce Marine AS; Per Erik Dalen from Aalesund Knowledge Park AS and GCE BLUE Maritime. This research is partially supported by the project "GCE BLUE Maritime Big Data" funded by Innovation Norway and the project "AIS and Maritime Transport" (No. 4506002) funded by MarKom 2020.

REFRENCES

[1] D. Keim, G. Andrienko, J. Fekete, G. Carsten, G. Melan, D. Keim, G. Andrienko, J. Fekete, and G. Carsten, "Visual Analytics: Definition , Process , and Challenges," *Inf. Vis. - Human-Centered Issues Perspect.*, pp. 154–175, 2008.

[2] H. Wang, X. Zhuge, G. Strazdins, Z. Wei, G. Li, and H. Zhang, "Data Integration and Visualisation for Demanding Marine Operations," in *Proc. MTS/IEEE OCEANS 2016*, Shanghai, China, 2016.

[3] H. Wang, S. Fossen, F. Han, I. A. Hameed, and G. Li, "Towards Data-driven Identification and Analysis of Propeller Ventilation.," in *Proc. MTS/IEEE OCEANS 2016*, Shanghai, China, 2016.

[4] G. Li, H. Zhang, B. Kawan, H. Wang, O. Osen, and A. Styve, "Analysis and modeling of sensor data for ship motion prediction," in *Proc. MTS/IEEE OCEANS 2016*, Shanghai, China, 2016.

[5] C. Plaisant, "The challenge of information visualization evaluation," *Proc. Work. Conf. Adv. Vis. interfaces - AVI 2004*, pp. 109-116, 2004.

[6] C. Plaisant, "Information Visualization and the Challange of Universal Usability," *Explor. Geovisualization*, pp. 1-19, 2005.

[7] S. Carpendale, "Evaluating information visualizations," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 4950 LNCS, pp. 19-45, 2008.

[8] E. L. Koua, a. Maceachren, and M. J. Kraak, "Evaluating the usability of visualization methods in an exploratory geovisualization environment," *Int. J. Geogr. Inf. Sci.*, vol. 20, no. 4, pp. 425-448, 2006.

[9] J. J. Van Wijk, "The Value of Visualization," 1987.

[10] J. Scholtz, "Beyond usability: Evaluation aspects of visual analytic environments," *IEEE Symp. Vis. Anal. Sci. Technol. 2006, VAST 2006*, pp. 145-150, 2006.

[11] J. J. Thomas and K. a. Cook, "Visualization Viewpoints: A Visual Analytics Agenda," *IEEE Comput. Graph. Appl.*, vol. 26, no. February, pp. 10-13, 2006.

[12] J. Thomas and J. Kielman, "Challenges for visual analytics," *Inf. Vis.*, vol. 8, no. 4, pp. 309-314, 2009.

[13] J. Stasko, "Evaluating information visualizations: Issues and opportunities," *Proc. 2006 AVI Workshop on Beyond Time and Errors: Novel Evaluation Methods for Information Visualization - BELIV'06*, 2006.

[14] C. A. Sharp, "Qualitative research and evaluation methods (3rd ed.) [Book Review]," *Eval. J. Australas.*, vol. 3, no. 2, 2003.

[15] T. D. Zuk, "Visualizing Uncertainty," *CGIM 08 Proc. Tenth IASTED Int. Conf. Comput. Graph. Imaging*, pp. 1–336, 2008.

[16] A. Tang, M. Tory, B. Po, P. Neumann, and S. Carpendale, "Collaborative coupling over tabletop displays," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '06*, p. 1181, 2006.

[17] M. Duffy, "Sensemaking: A collaborative inquiry approach to 'Doing' learning," *Qual. Rep.*, vol. 2, no. 2, pp. 1–6, 1995.

[18] D. M. Russell, M. J. Stefii, P. Pirolli, and S. K. Card, "The cost structure of sensemaking," pp. 24–29, 1993.

[19] C. N. Ehler, "A Guide to Evaluating Marine Spatial Plans," *Isnare Com*, 2014.

[20] A. E. Cravens, "Explaining Enthusiasm for MarineMap :," no. July, pp. 1–6, 2010.

[21] M. Q. Patton, "Qualitative Analysis and Interpretation," *Qualitative Research and Evaluation Methods*. pp. 452–477, 2002.

[22] C. Marshall and G. B. Rossman, *Designing Qualitative Research*, 3rd ed. Thousand Oaks, CA: SAGE Publications, 1999.