



From ethnographic research to big data analytics - A case of maritime energy-efficiency optimization

Downloaded from: <https://research.chalmers.se>, 2026-04-04 14:52 UTC

Citation for the original published paper (version of record):

Man, Y., Sturm, T., Lundh, M. et al (2020). From ethnographic research to big data analytics - A case of maritime energy-efficiency optimization. *Applied Sciences (Switzerland)*, 10(6). <http://dx.doi.org/10.3390/app10062134>

N.B. When citing this work, cite the original published paper.

Article

From Ethnographic Research to Big Data Analytics—A Case of Maritime Energy-Efficiency Optimization

Yemao Man ^{1,2,*} , Tobias Sturm ³, Monica Lundh ⁴ and Scott N. MacKinnon ⁴

¹ Department of Computer Science and Engineering, Chalmers University of Technology, 41702 Gothenburg, Sweden

² Department of Computer Science and Engineering, University of Gothenburg, 40530 Gothenburg, Sweden

³ Institute of Telematics, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany; tobias.sturm@student.kit.edu

⁴ Department of Mechanics and Maritime Sciences, Chalmers University of Technology, 41702 Gothenburg Sweden; monica.lundh@chalmers.se (M.L.); scottm@chalmers.se (S.N.M.)

* Correspondence: yemao.man@chalmers.se; Tel.: +46-(0)31-772-34-15

Received: 3 February 2020; Accepted: 15 March 2020; Published: 21 March 2020



Abstract: The shipping industry constantly strives to achieve efficient use of energy during sea voyages. Previous research that can take advantages of both ethnographic studies and big data analytics to understand factors contributing to fuel consumption and seek solutions to support decision making is rather scarce. This paper first employed ethnographic research regarding the use of a commercially available fuel-monitoring system. This was to contextualize the real challenges on ships and informed the need of taking a big data approach to achieve energy efficiency (EE). Then this study constructed two machine-learning models based on the recorded voyage data of five different ferries over a one-year period. The evaluation showed that the models generalize well on different training data sets and model outputs indicated a potential for better performance than the existing commercial EE system. How this predictive-analytical approach could potentially impact the design of decision support navigational systems and management practices was also discussed. It is hoped that this interdisciplinary research could provide some enlightenment for a richer methodological framework in future maritime energy research.

Keywords: ethnography; thick data; big data; machine learning; maritime energy efficiency; interface design; decision support; energy management; knowledge development

1. Introduction

As a critical transportation sector which accounts for estimated 80% of the volume of world trade [1], the shipping industry has been striving to reduce the environmental footprint in recognition of climate change challenges. In April 2018, International Maritime Organization (IMO) adopted an Initial IMO Strategy on reduction of Greenhouse Gas (GHG) from ships, aiming that we should reduce the total annual GHG emissions by at least 50% by 2050 compared to 2008 and further strengthen the energy-efficiency (EE) design requirements for ships [2]. EE in shipping is defined as energy used per transported goods and distance, e.g., kg of fuel per tonne cargo and nautical mile [3]. It is a multi-faceted issue and the ability of a vessel to decrease GHG emissions is a coordinated effort between speed control, navigational decisions, engine maintenance, hull resistance, propeller efficiency, scrubber systems etc. However, for the purposes of this paper, it is simplified as describing this as a sub-system which is directed at examining how speed governance can be regulated and predicted using machine-learning techniques. The intensification of EE, i.e., decreased fuel consumption while

maintaining at least the same level of transportation services, is considered as a significant contributor as it could lead to reductions of 25% to 75% of CO₂ emission, according to an IMO GHG study [4]. In addition to the environmental concerns, decreases in fuel consumption are also important to the economic sustainability of shipping companies [5].

There are multiple factors that could impact EE revealed by previous maritime research, such as optimizing weather routing (e.g., reducing resistance considering the wind, waves and current), minimizing rudder usage with adaptive autopilot settings, optimizing quantity of ballast water carried and optimizing trim and draft for lowest hull resistance [6,7]. Some research suggests that the most significant saving comes from voyage speed optimization—by lowering the speed by 10% the vessel can save approximately 20% in propulsion fuel consumption [6].

It has also been observed that the ship's crew usually have a direct and considerable impact on EE through their operational practices [3,8,9]. For example, as punctuality is critical for ferry vessels, the crew need to find the proper balance between fuel consumption at lower speeds and the direct and indirect costs of delayed port arrival [3]. However, insufficient knowledge and awareness of the practitioners and management are commonly identified to be one of the most critical gaps in making energy-saving decisions [5,7,10–13]. The challenges in the maritime energy field are characterized by elements such as compromised quality and low level of awareness of information [13,14], inadequate data analysis and actionabilities of improved operational measures [15], under-developed management standards [16], information asymmetries and power structures within organizations [17]. In the maritime organizational context these gaps are usually multiplied as the today's shipping industry essentially operates under a top-down management system [11] in which the practitioners' voices were not typically considered in ship–shore communications [18]. All these considerations constitute the complexity in the EE domain. But what methodological tools do researchers possess to understand the complexity existing within the domain?

The application of ethnographic research can contribute to a better understanding of EE practices, however, these approaches have been largely ignored within a technology-driven industry. In the maritime EE domain, an important advantage for taking such an epistemological approach is to move researchers closer to the tacit knowledge, norms, understandings and assumptions that are considered by the ship's crew during their work and are influential in forming the basis for decisions and practices that support systems design [19]. The importance of applied ethnography is frequently underestimated within a sociotechnical perspective. One example is that speed is usually considered as an important factor for EE which is relevant to decreased time to port. However, with in-depth interviews, it was revealed that ship–shore–port communication, time constraints for ship operators, means for predicting energy use, consequences of the breach of the rules of punctuality, as examples, all play a role in this complex sociotechnical system [3]. Sampson and Poulsen's recent study showed that cargo-owners' decisions can increase emissions because the speed choice for tankers can be strongly limited by cargo-owner's commercial considerations [20]. Many other studies have also shown that EE operation of the vessels can become irregular and complex as the picture is influenced by multiple actors and their relationships. There could be involvement of multiple stakeholders [21], goal or demand conflicts with EE operations (thus limiting the “maneuverability” for officers on board) [22], gaps in inter-department collaborations [23,24], social barriers in communities of practice [12,25], lack of trust between ship and shore organizations (thus downgrading the organizational practice) [26], etc. These insights in the EE domain can be captured by ethnographic research or its associated “thick data” approaches [27] (i.e., “a sticky stuff that is difficult to quantify” but “offers incredible depth of meanings and stories” [28]). It is important perceive and understand these ethnographic inquiries within the context of sociotechnical systems, so that we may make sense of the needs, requirements and even obtain inspirations for system design.

Unfortunately, such ethnographic inquiries and efforts usually stop before going to the nitty-gritty world of technical design and implementation for decision support systems, leaving the notorious issue of “implications for design” [29] and creating a “research-practice gap” [30]. This is essentially

making the designers and engineers of technical systems used onboard the goalkeepers, as they seem to be the only ones that can influence the design and use of the technology at the “sharp end”. In fact technology is but one essential part of the sociotechnical system [31] and it is vital that the design of the technological artefacts should consider these ethnographic inquiries. While EE performance monitoring, evaluation, prediction and decision support for improved operational and management practices is becoming increasingly important, it is also necessary to realize that this process of data collection and analysis is becoming increasingly difficult due to lack of appropriate energy consumption monitoring means and practices [9,15]. In the past several decades, many technology-centered solutions have thus been introduced to the maritime domain for better data analysis and management practices. In particular, some state-of-the-art machine-learning based approaches [32–36] were proposed to take advantage of big data in the energy field to improve EE and provide better decision support. The pace of this introduction is even accelerating in the wave of digitalization and full automation [37]. The increasing volume and diversity of ship energy consumption-related data are challenging the shipping industry to face a new landscape [38], entering a new era where data becomes the new oil and big data analytics becomes the means to generate novel services that can allow better performance evaluation, prediction and decision-making support [39–41]. There are some practical issues that ethnographic research cannot deliver alone without diving into the sea of big data analytics [35], which is an increasingly popular and important tool for tackling complexity in shipping.

However, this does not mean that big data analytics is the panacea. Understanding the EE practices and barriers in shipping is crucial [42] and it requires the work of ethnography or ethnomethodology that are essentially concerned about learning “tacit knowledge” and everyday social practice in the field [43]. There is a complementary relationship between ethnography and big data [44]. Unfortunately, almost all these state-of-the-art big data models were developed as a result of “decontextualization” by adhering to the physical law—they never (or never tried) associated with the disciplines of ethnography or ethnomethodology. For instance, it is extremely rare to see an interdisciplinary work in the maritime EE domain which machine-learning engineers cite ethnographic research work that allows big data analytics to be inspired by the thick data. Typical machine learning-oriented studies in the EE domain [32,33] did not really ground their ideas and feature selection in the model’s training phase on the insights produced by ethnographic research. This approach may raise serious issues that implemented technical solutions would never become truly useful [45] as the users’ real needs, experienced pains and use-of-context were “decontextualized” in the first place.

The review of literature above suggests that research that takes an interdisciplinary approach to understand operational EE gaps, as well as inform the design of a navigational decision support system, is rather scarce. Would it be possible to have an interdisciplinary research practice that uses ethnographic research to inform big data analytics (or at least inspire the need) and decision support system design in the field of maritime energy efficiency? Could we have a research practice that covers a holistic process of design, from the exploration in the field to technical implementation in the lab? What could be the value of a pluralistic epistemology embedded in an interdisciplinary research? This paper strives to provide insight into these questions by taking an exploratory approach to examine a case regarding maritime energy-efficiency optimization.

The purpose of this paper is to provide a paradigm to understand how applied ethnographic research could inform big data analytics (or at least the need) for designing a better decision support system in maritime energy context. It also aims to shape a promising space for future maritime energy research with a richer methodological framework. The following section describes the overarching methodological framework that is referenced to base the authors’ choices of methods at different stages and chart the exploratory path in a specific maritime EE context.

2. Methods

To better conduct this interdisciplinary research, the authors of the papers comprise ethnographic researchers with essential knowledge in machine learning and technical engineers with deep insights

in ethnographic studies and maritime operations, who have been working together in a collaborative manner since the start of the research project. The well-known “Double Diamond” design process and innovation framework, which was adapted from the divergence-convergence model proposed in 1996 by Banathy [46] and later popularized by the British Design Council starting in 2004 [47], was employed to guide this interdisciplinary study, undertaking an ethnographic study on ferry vessels and applying big data analytics for EE optimization (see Figure 1).

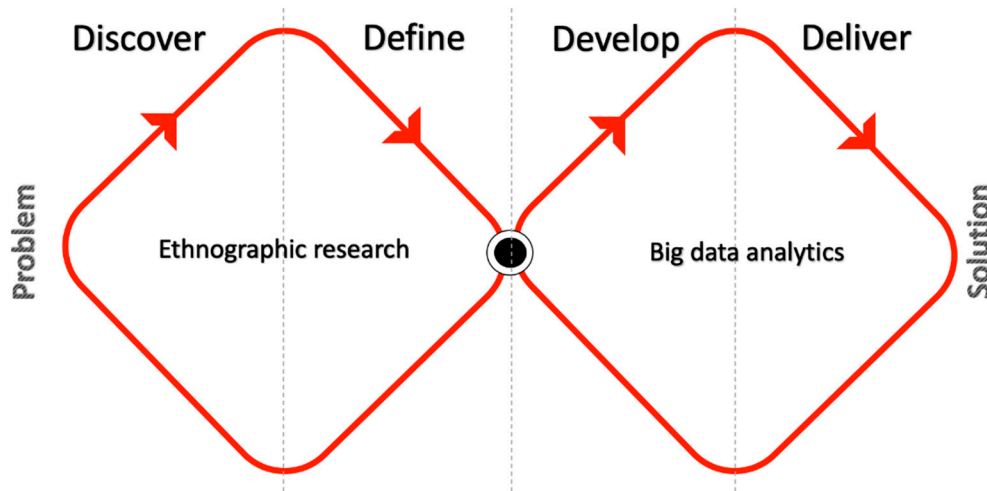


Figure 1. The overarching methodological framework adapted from the British Design Council [47].

An ethnographic study regarding the actual use of an advanced technological artefact used on ferry vessels (a commercial EE monitoring system) is positioned in the first diamond to understand a real problem onboard ship. This problem was co-investigated with Viktorelius who studied the same ferry vessels [8,12]. Ethnographic research is deemed as a qualitative method to understand human behaviors in their everyday practices [48]. Essential ethnographic methods include participatory observations and interviews, trying to discover the meanings behind the actors’ behaviors and beliefs [49]. Although this form of qualitative research may generate small data sets in quantity, the data is “thick” and delivers in-depth insights about the observed patterns and phenomenon [28,44]. Because doing ethnographic research means that the researchers need to be immersed in the fieldwork and see the things through the lenses of the field workers [43], it enables the research team to *discover* insights into the problem onboard and *define* areas to focus on, e.g., make sense of what is happening, how the crew members are using the technology or alter the way it is used, what detailed problems emerge in the whole interaction process, what is the underlying tacit knowledge or the taken-for-granted realities of everyday working life in navigation and energy saving, whether there is a need or opportunities for big data analytics onboard ships in the area of EE and if so, how the ethnographic study can inform the use of big data analytics to optimize EE operations and decision support system design. The answers to these questions will have a significant impact in consideration of the second diamond, even potentially the methodological choices. From this perspective, this holistic research paradigm in this paper is rather exploratory.

Big data analytics were initially situated in the second diamond as a potentially useful method to cope with the specific problem that were identified and defined in the first diamond. The research team needs to review the ethnographic study’s findings and request relevant data from the shipping company in order to *develop* corresponding machine learning-based solutions that may have the potential to address the observed issue (or at least make an improvement of the current situation). The findings of the experiments will also be *delivered* to the shipping company as the basis for future research and development.

3. An Ethnographic Study on Ferry Vessels

To explore the practical energy-saving issues in the use of advanced technologies, the authors of this paper planned an ethnographic study and visited the ferry ships owned by a collaborating ferry company. The company introduced a commercial fuel consumption monitoring system called ETA-pilot to assist the navigators to regulate the speed automatically, as speed is directly related to fuel consumption. Voyages had been divided into multiple legs with fixed positions (waypoints). The ETA-pilot proposed an optimal speed setting for each leg of the journey, based on multiple factors (e.g., ship trim/draft, depth of the water, weather information, distance to the destination and estimated time of arrival, etc.). The speed can also be adjusted by the navigators as course and speed deviations may be required to follow collision avoidance regulations. The fuel consumption (per nautical mile) is displayed as a dynamic curve along with other output parameters in a complex line chart at the bottom of the user interface, while the total consumption is displayed on the top right corner (see Figure 2).



Figure 2. The true speed (17 knots/h), the fuel burning rate (123 kg/nautical mile), as well as total fuel consumption (10,696 kg) were displayed on the user interface of the ETA-pilot.

In ethnographic study fieldwork, all research team members were learning to become functioning members of the practitioners' community of practice through introspection and intersubjective inquiry in order to better "percolate" the tacit knowledge to the surface [43]. The authors actively observed how the deck and engineering officers worked in situ, documented the context of use and their purpose of use in their everyday work and interviewed them to understand their perceptions of how this seemingly advanced technology influenced their knowledge and awareness in saving fuel. Documentation from observations, interviews and other sources of information were sought and used because no single source of information could be trusted to provide a comprehensive perspective on the whole user experience [43]. Using a combination of these data collection approaches enabled the research team to validate and crosscheck findings.

What was observed as a significant finding was that the bridge crew frequently disabled the ETA-pilot and manually set the speed and course based on their tacit knowledge around the condition of weather and dynamic traffic situation, claiming that they could improve upon fuel consumption. It was observed that they were frequently monitoring speed, course and wind situation. A higher speed is usually preferred as the commercial risk of a delay is very high, considering the requirements on punctuality, but they were also aware that a higher speed could lead to more fuel consumption, especially in shallow water and head wind. The crew did not base their decision on whether to

deactivate the ETA-pilot or which speed to set on formulas, but instead, they relied heavily on traditional navigational instruments and their own navigational experience to maneuver the ship. Much of the voyage and post-voyage information provided by the ETA-pilot was not observed being used by the crew members. On one occasion the captain was actively introducing the features of the ETA-pilot to the research team, but because he realized he had received some wrong information about the weather (which was required to be put into the machine for it to function properly), he had to turn off the ETA-pilot completely before departure.

That enthusiasm of using the tool was not common amongst the crew members. Instead many of them pointed out that the underuse of the toolbox was subject to a lack of trust and usable information—it was difficult to know if the tool was indeed helpful for reducing the fuel consumption or not as there was no benchmarks for comparison, let alone decision support. A fancy line chart was plotted on the display but almost no seafarers used and reviewed it during or after a journey. A tool may be functionally powerful but if it does not address the true needs or failed to be integrated into the work practice, it will be risking becoming an ornamental bauble. During the voyage, the ETA-pilot basically only presented two easily readable information sources that the navigators could instantly monitor without much cognitive effort: the consumption rate and total fuel consumption. But according to the navigators, such information can hardly be used for real-time self-evaluation and navigational decision support in an eco-driving manner, which was explicitly mentioned multiple times during the interviews as their “wish-list”. The tool did not really provide the users a learning space for knowledge and awareness improvement. Several navigators commented that if the ETA-pilot could see some sort of fuel consumption prediction they might be able to be more proactive instead of being reactive in navigation choices. In addition, there could be many traffic situations in which the master mariner felt it was necessary to increase the maneuverability by increasing the speed, which could lead to more fuel consumption. Under these circumstances, they did not really mind the consequence of burning more fuel. To them safety always overrides EE goal. There appeared to be no incentives or information feedback to allow them to make more deliberate efforts to save fuel within their safety boundaries. All these technical configurations and dynamic situations revealed parts of their context of use. The use of the ETA-pilot was opportunistic at best. From the design and use of the tool’s perspective, a prominent issue began to surface: the operational choice for navigational strategies have a considerable impact on fuel consumption [8,13], but there is an absence of efficient monitoring in a real voyage environment because the navigators did not get sufficient analytical support from the existing EE monitoring tool.

This gap was not bound to the ship side. According to the supplier of the ETA-pilot system, the recorded real-time data on fuel consumption and weather condition were transferred to the shore-side for possible further analysis and hopefully EE optimization suggestions could be delivered back to the ship. However, the crew members expressed that they never received such feedback from the supplier except the overarching requirement to save fuel or taking the training course of the ETA-pilot on how to undertake eco-driving with its “decision support functionalities”. Once the vessel had reached her destination, that was the end of the role for this technology. Both the real-time decision support system and evaluating activities were missing. This phenomenon was also noted by Viktorelius who was involved in the ethnographic studies on the same vessels too, concluding that this was a problem of “underdevelopment of evaluating activities” or “inadequate knowledge” [12].

One of the authors of this paper also had a chance to interview the designer of the ETA-pilot. The designer explained that there was no universal optimal use of this decision support tool but this depended upon the navigator’s own experience and interpretation, as there are many factors that the current algorithm of the tool cannot universally account for, e.g., different routes, different efficiency from different propellers, different potentials to save fuel on different ships; the tool was not developed in a way that it can dynamically “learn” from the historical data and provide real-time decision support for eco-driving (although it does have some predictive capability such as adjusting speed based on

depth of the water and manually input weather information etc.). The designer admitted that such needs may be considered in its future development.

Overall, the lack of performance evaluation and decision support has contributed to the unexpected EE practice (advanced technologies being not used most of the time) and potential issues in reducing fuel consumption, whereas the shipping company/developer of the tool seemed to lack the competence or resources to analyze the large amount of data that was generated during voyages. The ethnographic study did not only help to identify some particular parameters of concern (speed, course, wind, fuel consumption), but also identified some actual gaps in the use of a “black-box” technology and needs of the users in their actual daily work for improved EE performance.

The most important insight is perhaps the associations made from the “first diamond” to the “second diamond” informed by the findings of the ethnographic studies. The idea and need of using big data analytics for better decision support for EE optimization in this context was much appreciated. It may introduce some changes that is important for the system improvement. The findings discovered that (at least) the current design and implementation of the tool had an intrinsic lack of capacity to enable improved operational practice regarding EE. This actual deficiency of the EE monitoring tool, although was merely a part of the sociotechnical factors, matched the strengths and advantages of artificial intelligence (AI) technologies and machine-learning techniques well. For example, the technologies will have the computational power to analyze data and support the learning and understanding of actual fuel consumption without adding extra operational workload upon ship personnel. Furthermore, the ethnographic research suggested some strategic direction that was worth further investigation, e.g., how to use historical ship sensor data to predict fuel consumption in a similar condition and provide easily understandable information to support decisions so that it may be possible for the ship operators to understand which factors are influencing the prediction, to what extent and how to adapt operational practice in situ. Additionally, the ethnographic research also elucidated some features that likely play significant roles in machine-learning modelling, such as speed, course, locations, wind, fuel consumption. Some studies used physics-based model simulations to estimate ship performance [50] and they usually did not integrate weather information into the modelling which will likely influence the prediction capability when used in a real voyage environment.

The purpose of the remainder of the paper is not to shift to the big data analytics to engineer a fully-fledged technological product as such, but to demonstrate a proof of concept that is inspired and initially informed by the ethnographic research, in order to complete this interdisciplinary approach, because the ethnographic research sets the departure point of the research on big data analytics and the follow-up discussion.

4. Applying Big Data Analytics for Maritime Energy Efficiency

The research activities in the second diamond were to use machine learning to predict fuel consumption given an operational and environment profile. In order to do so, it is necessary to acquire enough data from various sources. The research team requested and obtained voyage data recorded over a one-year period from five different European-based merchant ferries that are available to carry passengers, cars and trucks. In this section of the paper, we will walk into the technical world and propose to construct two models of the propulsion systems of the ships by using supervised and unsupervised machine-learning techniques and evaluated their performance in terms of data prediction accuracy and generalization capabilities. It is followed by a discussion around the potential impact of these predictive-analytical approaches on the interface design and management practices optimization.

4.1. Data Sources and Preprocessing

Sensors collect data continuously on the ships, but only store an averaged value every 10 min. Relevant sensor data sets based on the findings of ethnographic research, e.g., speed over ground (SOG), speed through water (STW), apparent wind speed (AWS), course over ground (COG), fuel consumption, etc., from five different vessels (Ship 1, 2, 3, 4, 5) of the same shipping company and

automatic identification system (AIS) data were also obtained. Most attention was paid to the Ship 1's journey between Gothenburg and Kiel, which is longest in distance among all the voyages run by these five vessels. For vessel 1 and 5, the data covers approximately a two-year period and just a one-year period for the remaining vessels. To clarify how the data were pre-processed, the following concepts are provided:

1. A data point is defined as a 1-dimensional value that represents a 10-min average, starting at a defined point in time. A complete list of data points collected by the on-board ship sensors is provided in Table A1 in the Appendix A. Data points can also carry static values (such as ship identifier) and may not relate to an on-board sensor.
2. A measurement is defined as all data points that belong to the same starting point in time.
3. A data row (or row) is a measurement, extended with values that can be derived from measurements or originate from other data sources than on-board ship sensors.

Data were provided in Microsoft Excel files (1 file per ship). Columns were renamed or moved to ensure that all files have the same set and order of columns. All the string values are also converted to numerical values. Measurements deliver a continuous stream of information; thus, it is more practicable to divide this stream into sections (separation into journeys). A journey index was assigned to rows of the sensor data. An assumption was that a speed greater than 8 knots marks the beginning or continuation of a journey and a speed below 5 knots marks the end of a journey. The data from different sources had to be integrated, assimilated and normalized and are further described in the following sections. Figure 3 presents an overview of the data pipeline in the pre-processing stage.

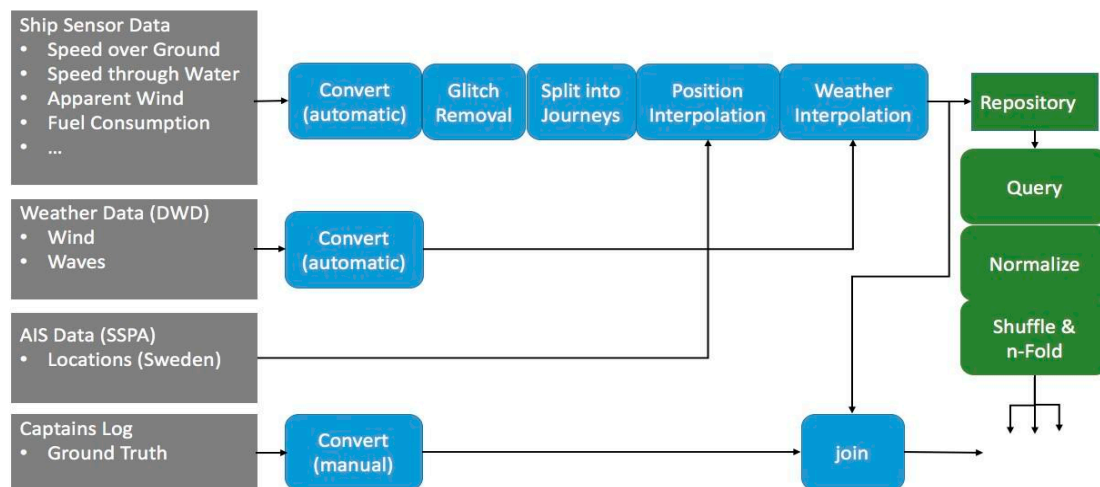


Figure 3. Data pipeline to provide an overview in data pre-processing.

4.1.1. Data Assimilation

Weather data was acquired once for all journeys, before the machine model was trained, in order to speed up the training process. Historical weather forecasts data were obtained from forecast.io (An online service which aggregates multiple weather models—see forecast.io/raw) and another data set was provided by the German Meteorological Office (DWD) (Deutscher Wetterdienstsee https://www.dwd.de/DE/Home/home_node.html) Data points included in the sets are shown in Table 1. Weather forecasts describe values for rectangular grid cells at defined time intervals. For the DWD data, grid cells span 0.05° in south-west and 0.1° in east-west (equivalent to 5.57×5.90 km/ 5.57×6.54 km) and are generated for 3, 6, 9, and 12 h in one model run. Model runs were performed every 12 h. Interpolation of time and location was used to retrieve weather information from the DWD data set.

Table 1. Weather data sets.

Data Point	forecast.io	DWD
wind speed	X	X
wind bearing	X	X
pressure	X	
temperature	X	
humidity	X	
signi. height waves combined		X
signi. height wind waves		X
mean direction wind waves		X
direction swell waves		X

Ships and traffic control use the broadcasted information from an automatic identification system (AIS) to monitor other vessels' navigational intentions. The AIS data provided by SSPA (A Swedish company which provides maritime solutions—see <https://www.sspa.se>) covers information about ships up to 60 km from the Swedish coastline. This provides information on the direction of travel for each journey, checking whether the vessel was located in a 10 km radius of the pier of Gothenburg at journey start time.

Unfortunately, the AIS data provided cannot be used to reliably track a vessel during a complete voyage, it is necessary to get location approximation by using speed. All average speed (km/hour) samples over all 10-min intervals of the current journey (starting at t_0), divided by 6 results in traveled distance relative to journey start (see Equation (1)). Taking the voyage from Gothenburg to Kiel as an example, the position can be approximated using a linear interpolation via relative distance travelled. The track used consists of 83 waypoints that span a total length of 437.43 km.

$$\text{dist}(t) = \sum_{i=t_0}^t \frac{\text{speed}(i) * 10}{60} \quad (1)$$

4.1.2. Removal of Glitches

Visualization of onboard sensor data showed unexplained spikes that occur in single data points. Various reasons as malfunction of sensors, data transmission, data conversion, etc. could be the cause. Separation of journeys as described earlier also led to journeys that are too long (span over multiple actual journeys) or are too short (faulty speed measurement during voyage leads false detection of end-of-journey). Glitches as observed in the sensor data represented values far of the normal range of average data points. By counting all values of a data series, a threshold value could be determined so that all extreme glitches can be removed without losing too many data samples segment.

Figure 4 depicts the relationship between the top x%-values dropped and the corresponding amount of the remaining data points (i.e., the data that do not fall into the cut-off segment).

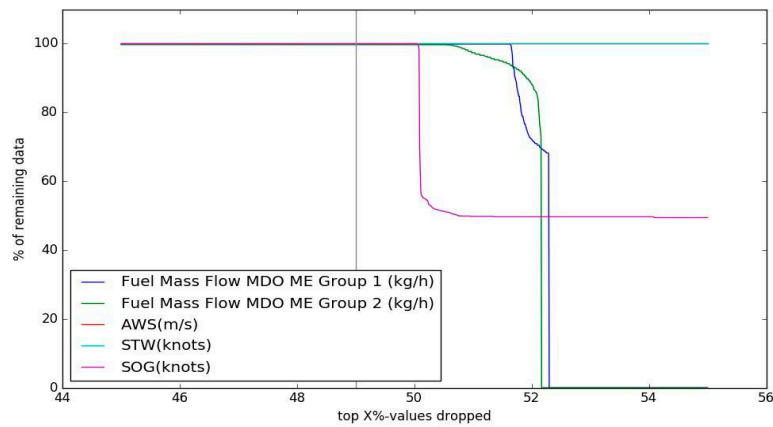


Figure 4. Number of remaining samples per dimension after glitch cut-off.

As shown in Table 2, the number of cut-off data points by a threshold of 47% is fairly low and introduces minimal loss of data.

Table 2. Amount of remaining data points per series after cut-off at 47%.

Dimension	Total	Removed	Remaining in %
Marine Diesel Oil (MDO) Main Engine (ME) Group 1	50,958	44	99.912
Marine Diesel Oil (MDO) Main Engine (ME) Group 2	50,958	144	99.716
Apparent Wind Speed (AWS) (m/s)	50,767	35	99.993
Speed Through Water (STW) (knots)	45,419	31	99.993
Speed Over Ground (SOG) (knots)	50,785	20	99.996

4.1.3. Normalization

Normalization is always performed on the same data set that is used for model training. Z-normalization is used in this work to ensure that all elements of the input vector are transformed into the output vector whose mean is approximately 0 while the standard deviation is in a range close to 1 [51]. Before normalization, data will be filtered to remove all entries where fuel consumption of all fuel types is recorded as 0, to avoid a 0-bias in the average. In the example shown here, 26,925 data rows were processed (from a total of 248,128). Table 3 shows exemplary three dimensions before and after the normalization.

Table 3. Exemplary three dimensions data normalization.

Dimension	Min	Max	Mean	Std
AWS (m/s) normalized	0.000000 -1.918170	43.038380 4.194465	13.505619 0.000000	7.040889 1.000000
COG (degrees) normalized	0.000000 -1.490091	359.990000 1.791616	163.456975 0.000000	109.695950 1.000000
SOG (knots) normalized	0.000000 -3.137029	21.129949 1.135663	15.513702 0.000000	4.945348 1.000000

4.2. Constructing Prediction Models

Two approaches on how machine learning is used to process the data described in Section 4.1 are considered, in order to provide a predictive-analytical foundation for future energy-efficiency monitoring tools. The selection of the input and output parameters from these data sets are described in this section.

4.2.1. Multi-Layer Perception (MLP)

The multi-layer perception (MLP) is organized in layers, each layer consisting of multiple nodes (neurons). Each neuron has input connections and creates an output that is propagated to other nodes in the following layer. Two following layers are fully connected to each other. Weights are assigned to each connection ($w_{p,q}^{(l)}$ with $p = 0 \dots m^{(l-1)}$ and $q = 0 \dots m^{(l)}$). The output depends on the weighted inputs and an activation function. The number of nodes and the used activation function can vary between layers. For convenience the 0th layer is named *input layer*, the $(L + 1)$ th *output layer*, and all other layers are *hidden layers*. The number of nodes in each layer ($m^{(l)} + 1$) is called the dimension of layer l . Values set in the input layer are typically described as input vector $x = (x_0 \dots x_D)^T$ and values of the output layer as vector $\hat{y} = (y_0 \dots y_C)$. Weights for layers are described via matrices. The dimension of the input layer is determined by the number of input dimensions from the data sets, see Equation (2). A classic multi-layer perceptron introduced by Rosenblatt [52] is described in Figure 5.

$$y_q^{(l)} = a \left(\sum_{i=0}^p w_{i,q}^{(l)} * y_i^{(l-1)} \right) \tag{2}$$

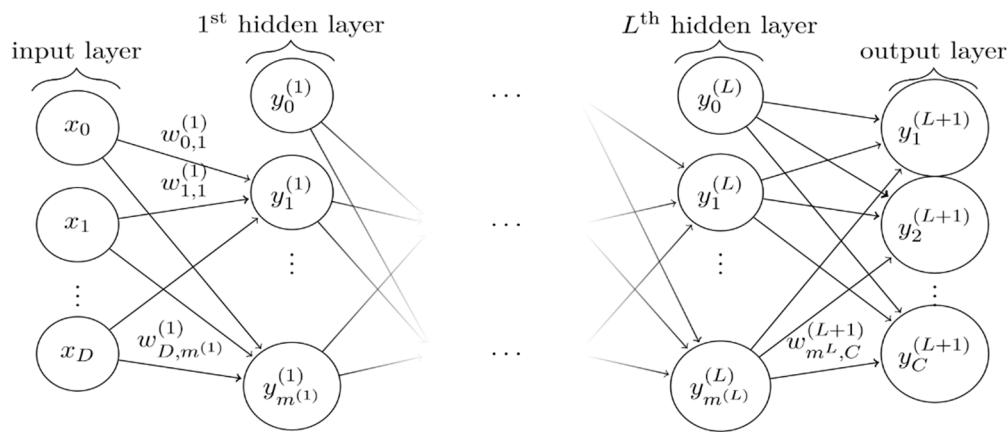


Figure 5. Multi-layer perception (MLP) with L hidden layers.

A data row can be split into an input and an output vector. For a model that predicts fuel consumption, Fuel Mass Flow Heavy Fuel Oil Main Engine (HFO ME) Group 1+2 and Fuel Mass Flow Maritime Diesel Oil (MDO) ME Group 1+2 are defined as y_0 and y_1 . A choice of parameters such as AWS knots, LWS (longitudinal water speed) knots, relative wind speed (m/s), and others are chosen for x (see Table A1 in the Appendix A for explanation of parameter abbreviations). To use the model for predictions, x is set and the network will compute \hat{y} using all $w^{(L)}$. Cost (\hat{y}, y) defines a cost function that measures the difference between expected output y and calculated output \hat{y} (prediction). Training the model means to find a configuration for all weight matrices $w^{(L)}$ such that the sum of costs for all predictions over the set of training inputs is minimal. Training is performed using back-propagation—an algorithm in which the error of a prediction is propagated backwards into the weight matrices, depending on the weight’s contribution to the error as described by Rumelhart et al. [53].

The initial hyper parameters have to be chosen before the model is trained. In this paper hyper parameters are:

- activation function (logistic, arctan, softsign, rectifier, softplus, gaussian, ...);
- learning rate (initial value, time decay);
- learning batch size;
- cost function;
- number of hidden layers;

- size of hidden layer;
- choice of input parameters.

This opens a rather large search space to perform a brute-force search. Through a few benchmarks, an activation function was chosen to be the sigmoid function for all layers, except the output layer, for which we chose the identity function. Our cost function is the MSE (mean-squared error) between normalized predictions and normalized expected outputs. The learning rate is set to an initial value of 0.01 and then determined by the Adam algorithm [54]. Batch sizes were initially tested for values 1, 5, 15, 50, 100, 250, and 450. For further tests 50, 150, and 450 were used. To determine the set of input parameters, number of hidden layers and their dimension we followed the flow-chart depicted in Figure 6.

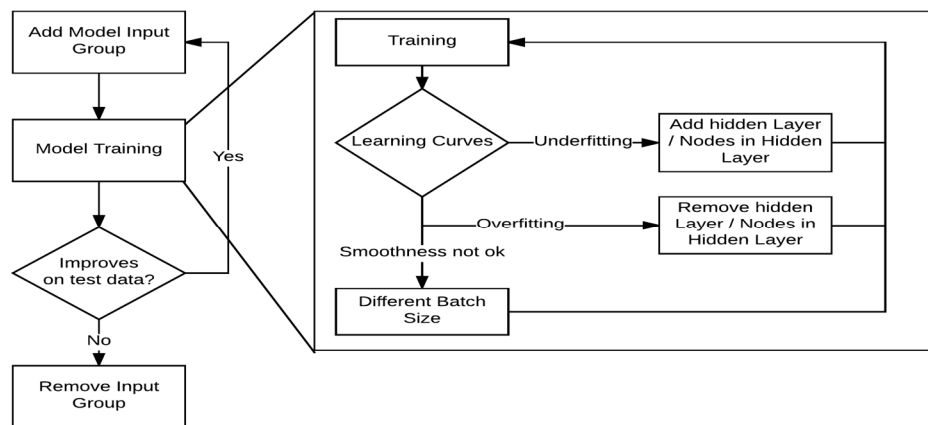


Figure 6. Search for MLP hyper parameters.

The six different sets were evaluated. Their input parameters are grouped into six groups, from A to F. They were listed in Table 4.

Table 4. Input parameter groups (A to F).

Data Point	A	B	C	D	E	F
Direction	X	X	X	X	X	X
Apparent wind angle (degrees)	X	X				
Apparent wind speed (m/s)	X	X				
Current direction (degrees)	X	X	X	X	X	X
Course over ground (degrees)	X	X	X	X	X	X
Water depth (m)	X	X				
Speed over ground (knots)	X	X	X	X	X	X
Speed through water (knots)	X	X	X	X	X	X
Sea water temperature (°C)	X	X				
True wind direction (degrees)	X		X	X	X	
True wind speed (knots)	X		X	X	X	
journey-time-traveled	X	X	X	X	X	X
wind-speed(m/s)-dwd	X	X	X	X		
mean-wind-wave-direction(degree)-dwd	X	X	X			
swell-wave-height(m)-dwd	X	X	X			
mean-swell-wave-direction(degree)	X	X				
wind-wave-height(m)-dwd	X		X			
wind-direction-(degree)-dwd	X		X	X		

Figure 7 depicts the MSE between predictions of MLP models and validation data, while testing different models with different parameters. The MSE values are the average over 5 runs per model configuration (5-fold). Each configuration was trained with batch sizes of 50, 150 and 450. One can see,

that with more input parameters, some models could not be trained and resulted in MSE values close to or greater to 1. Additionally, models with larger parameter number tend to converge faster than models with fewer parameters (shorter training time, because training is aborted when the model does not improve).

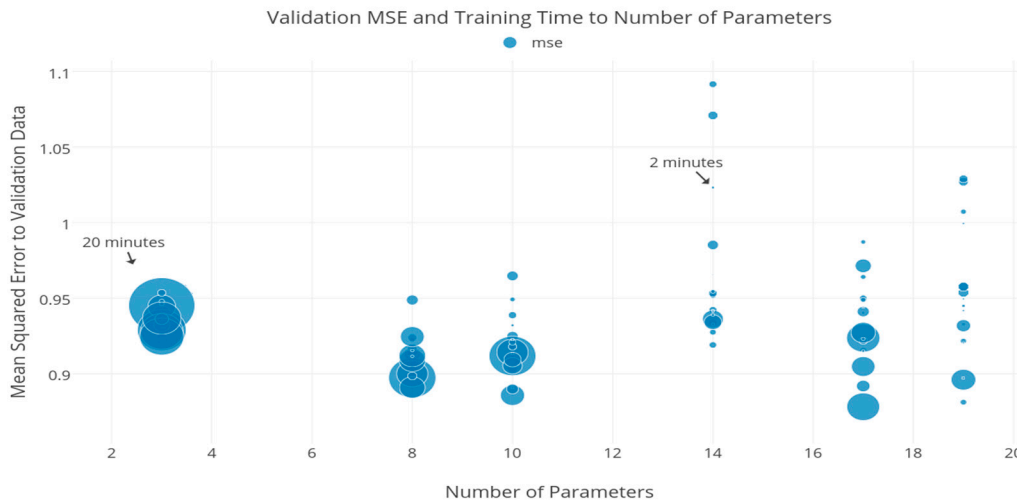


Figure 7. Validation mean-squared error (MSE) and training time of MLP models with depending on the number of parameters.

Figure 8 shows the relative error to the validation data for different choices of input parameters. This relative error is the average over all model runs with the described input parameters (varying network size, batch size, etc.). The selection of some parameters depended on the inclusion of other parameters, because they are connected semantically (e.g., wind speed and wind direction). It is clearly visible in Figure 8 that too many input parameters led to models that were not trainable. Given the amount of data captured (1 or 2 years per ship), the number of trainable weights is too high for the MLP to converge.

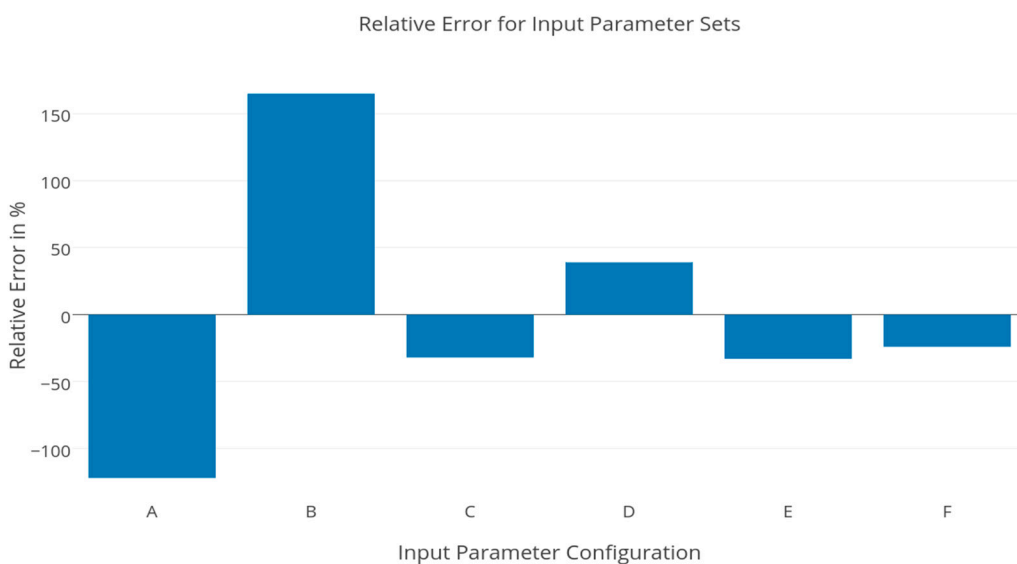


Figure 8. Relative error on validation data for input parameter groups.

After intensive search for the best combination of hyper parameters, two MLPs, each with one hidden layer, containing 10 neurons, were chosen to be the best model in combination with parameter set E (8 parameters, no weather) or D (10 parameters, including wind data). A batch size of 150 was

found as the optimal for this data set, when used with an Adam optimizer [54]. Figure 9 shows the improvement of the MSE over the test data set over training time (number of samples seen) of the model with parameter set E.

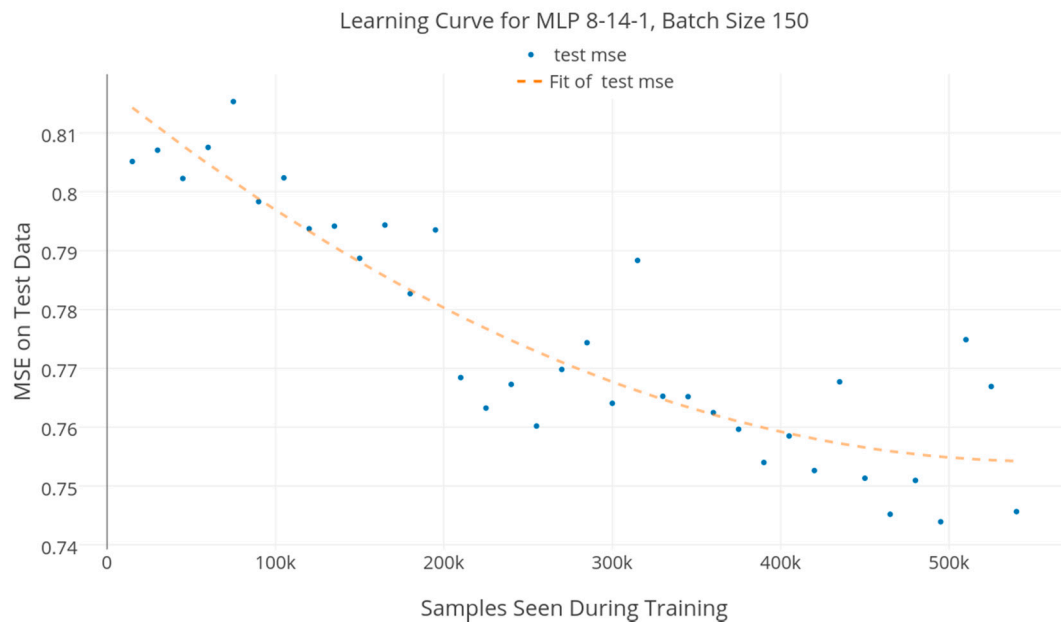


Figure 9. Learning curve of final MLP model.

4.2.2. Self-Organizing Map (SOM)

Self-organizing maps (SOMs) can be trained with unsupervised learning, which allows us to derive structure from data where we don't necessarily know the effect of the variables. The algorithm is designed to find correlations between inputs and outputs using competitive learning (instead of error-correcting learning like back-propagation). A high-dimensional input space is reduced to a typically two-dimensional discrete space. Due to SOMs' non-linear properties, they can be seen as non-linear biologically inspired generalization of principal component analysis [55,56].

Continuous high-dimensional space is mapped to a low-dimensional and discrete space. Human brains are known to be dividable into different regions that are more active for certain tasks like hearing, seeing or motion. Neurons in a SOM are located in the same way, as soon as the SOM has converged.

The input for a SOM is a vector that correlates to $(x_0, x_1, \dots, x_D, y_0, \dots, y_C)$ as described in Section 4.2.1. Neurons are organized in an n -dimensional space (typically $n = 2$), called lattice s . Input nodes are fully connected with all neurons in the lattice. One weight $w_i = w_{s,i}$ (or synaptic connection) is assigned to each neuron i and initialized with a small random number. Like MLP, the SOM has a training and a mapping phase. During training, three steps are performed:

1. Competition: each neuron calculates its output via $w_i^T x$ depending on input vector x . The neuron with the largest output value is the best matching unit (BMU). If the BMU is excited, it means that its weights will be updated (synaptic adaptation).
2. Cooperation: the neuron's location in the lattice implies a neighboring topology to other neurons in the lattice. Neighboring neurons will also be excited and also undergo synaptic adaptation.
3. Synaptic adaptation: excited neurons change their synaptic connections to increase their output value for the given stimuli. The degree of change depends on the distance to the BMU, whereas the BMU will perform the most severe synaptic adaptation among all excited neurons.

After training, the neurons represent a feature map. Their lattice location relative to each other represents properties from input vectors in the higher dimensional input space. A converged SOM meets three properties:

1. The feature map can be used as a decoder and encoder where competition mechanism (finding the best matching unit) is equivalent to $c = \text{encode}(x)$. The weight vector w_c represents the decoded value $x' = \text{decode}(c)$.
2. The SOM is topologically ordered according to features of the input space.
3. When region on the input space occurs more often during training, more neurons will represent the mapping for these inputs in the feature map (the feature map will have a high resolution for these regions).

In this paper, SOM is used to make predictions for fuel consumption and find a similarity order among journeys. If the SOM converges, it can be used to predict fuel consumption for vessel situations (data rows, or input vectors) that are not included in the training data, if there are enough training data vectors that are similar to the unseen input vector. Therefore, the process of finding a BMU has to be altered, such that the Euclidean distance can be masked for some entries of the input- and weight vectors. The SOM was trained on the input vectors with the dimensions speed through water, relative wind speed, draught, and fuel consumption can be queried for a prediction of fuel consumption by masking this entry on all weight vectors and the input vector when comparing them regarding Euclidean distance in the input space, to find the BMU. After the BMU is found, the previously masked entries are the result of the prediction query. This process is called encoding, because the high-dimensional input vector is mapped into the low-dimensional space:

$$\begin{aligned} \text{Enc}(x) &= \arg \min_i (\text{mdist}_M(x, W_i)) \\ \text{mdist}_M(x, W) &= \sqrt{\sum_{j \notin M} (x_j - x_i)^2} \end{aligned} \quad (3)$$

Encoded values can be compared using the lattice distance between neurons. We define the similarity distance between journeys as:

$$\text{jdist}(A, B) = \frac{\sum_{(a,b) \in A \times B} \text{mdist}_M(a, b)}{\|A \times B\|} \quad (4)$$

with A B set of joined data rows of re-sampled journey A and B . Two journeys may contain different number of data rows. To overcome this issue, all journeys were re-sampled before comparing them. Instead of using time as the data row index, we chose *relative distance travelled* and set the number of desired data rows to 100 (the longest route of all ships contains 84 data rows). Re-sampling is done by first calculating the relative distance travelled for each data row, using the sum of speed values of all previous data rows and the time since start of the journey, divided by the total distance of the journey. The results are attached as new column to the data rows of the journey, set as index, and then used for spline interpolation for index values 0 to 1 with 0.01 increments. The parameters for the SOM were chosen in the same way as hyper parameters of the MLP. After extensive tests, the parameter set D in combination with a lattice size of 20×20 neurons was found as the best-learning model.

In conclusion, we tested different input parameter sets in combination with different hyper-parameter configurations. All experiments have been conducted using the 5-fold validation method. A MLP model with up to 10 parameters can be reliably trained on the given data set. Batch sizes of 50, 150, and 450 have been evaluated and 150 has been found to give the best learning results. The topologies of $10 \times 10 \times 1$ and $8 \times 10 \times 1$ (parameter set D and E) were found to result in the lowest MSE on the validation data. More than one hidden layer could either not be trained reliably on the small data set or did not yield better MSE results. A SOM of a size of 20×20 neurons and parameter set D was found to have the lowest MSE on the validation set.

4.3. Evaluation

In this section, we present the evaluations of the two neural networks by comparing them to each other and to the ETA-pilot, which was described in our ethnographic study.

4.3.1. Comparing MLP versus SOM Predictions

We use two evaluation metrics to compare the two different models (SOM, MLP) with each other regarding accuracy and generalization and regarding accuracy to existing tools. All evaluations are performed 5-fold: models were trained on 80% of the data and evaluated on the remaining 20% (the validation fold). This process was repeated five times, such that every evaluation was included in the training data set four times and used for evaluation once each. Described evaluation metrics in this section are based on the average over these five runs. Models are evaluated on the following metrics.

1. Accuracy: relative difference between outputs and expected outputs.
2. Generalization: like accuracy, but for models trained on a subset of ships and validated on another, disjunct set of ships.

We evaluated the following model variants regarding accuracy and generalization for each of the model types (SOM and MLP) (Table 5). All models are the same regarding hyper parameter, except the dimension of the input layer for experiment A.

Table 5. Evaluation configurations.

Experiment	Training Data Set	Generalization Test on
A	Ship 1, with weather	Ship 5, with weather
B	Ship 1	Ship 5
C	Ship 1, Ship 5	Ship 2 and 4
D	All, except Ship 3	-

Ships 1 and 5 operate on the same route, whereas ship 2 and 4 operate on two further routes. Ship 3 was not used for evaluation because the data quality was not sufficient (too many missing values, glitches, etc.). The output of the model (prediction) is the estimated fuel consumption at the time of the given inputs included parameter groups D or E (see Table 4). The signed relative error between model output and measured fuel consumption is computed. We used a signed metric for these experiments to see if models were biased to compute results that are too large or too small.

Figure 10 shows the distribution over all evaluated samples. Boxes mark second quartile, median and third quartile. Whiskers indicate first and fourth quartile. The line inside the box marks the median, the upper and lower edge of the box the second and third quartile. The whiskers around the box depict the range of the first and fourth quartile. The MLP model performs best in experiment C, when looking at accuracy. This could be explained through the increase of available training data compared to experiments A and B. Models of experiment A and B predict too low fuel consumptions on the generalization data set. Ship 5 might have different engines or is built in a way that increases the fuel consumption compared to Ship 1. Experiment C shows that a model that was trained for a specific route cannot be used to predict fuel consumption for other routes. Prediction errors range from -55% to $+170\%$. We can see that models that have been trained on a limited set of routes do not generalize well to predict fuel consumption for other routes. It is not acceptable to build reliable prediction tools on a mode with a relative error from up to 50%. More data have to be acquired to make these models more robust. SOMs do perform worse than the MLP models throughout all experiments. Relative errors of SOM experiments have a median close to 0% but spread to up to over 100%. Although SOMs are, like MLPs, not robust enough to be used as solid predictors, considering the data set we obtained (just over a one-year period), the chosen hyper parameters enable the model to be generalizable, when trained on different ships, meaning that the model's performance is not dependent on a single ship.

This is also manifested by the fact that the model that was trained on multiple routes and multiple vessel voyages perform as well as the models that were trained on only one or two vessels on the same route (experiment D).

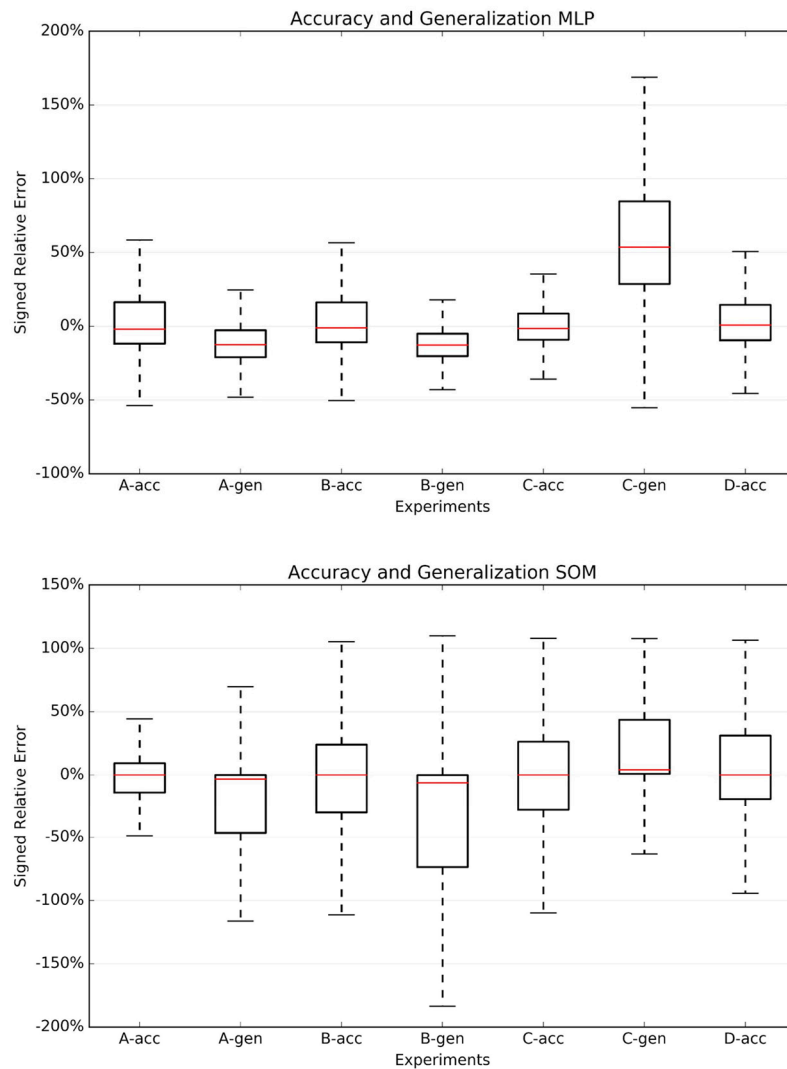


Figure 10. Accuracy (acc) and generalization (gen) of MLP and self-organizing map (SOM) models.

4.3.2. General Prediction Benchmark

We also compare our MLP and SOM models to the ETA-pilot tool. Ground truth for the experiments is the summed measured fuel consumption per journey for the Ship 1 in 2015. A total of 218 journeys could be matched to the captain’s log, which includes the estimation of the ETA-pilot that was issued before each journey, depending on the weather forecast. The average over summed relative error per journey for each model is shown in Table 6. One limitation we have is that there is no data specifying, whether journeys were performed using the recommendations of the ETA-pilot or not. The beginning and end of the voyages were not included in the experiments, because the ETA-pilot predictions do not include maneuvering close to port. The results of the models are based on the predicted fuel consumptions for every recorded data point for each journey. All predictions for one journey were summed and then compared to the expected fuel consumption for the given journey. The relative distribution of the difference to the expected fuel consumption is plotted (see Figure 11).

Table 6. Average over summed relative error per journey for each model.

Model	Average	Standard Deviation
Baseline (ETA-plot)	5.4	±17.8
MLP	1.4	±16.3
SOM	1.1	±20.5
MLP(Weather)	1.9	±15.8
SOM(Weather)	4.2	±21.4

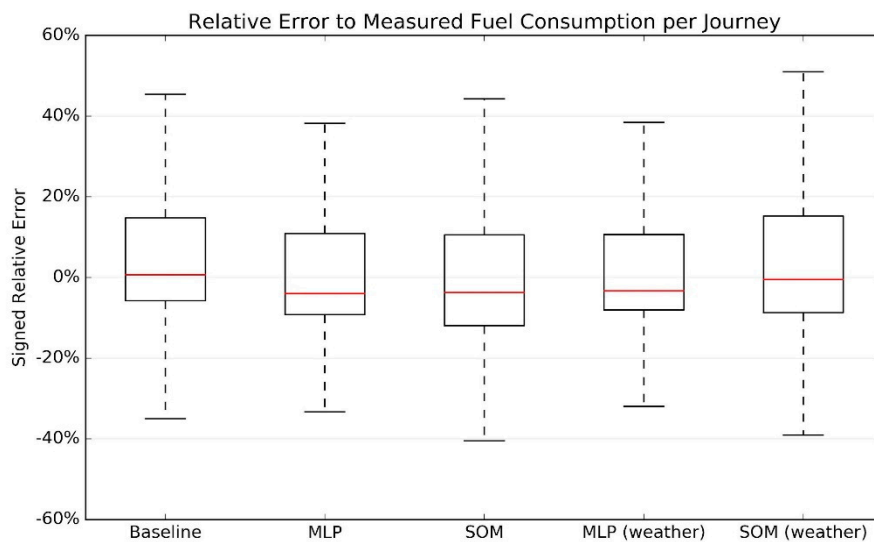


Figure 11. Relative error to measured fuel consumption per journey for baseline, MLP and SOM.

The average relative error and the standard distribution is calculated from the same result set as the box plots in Figure 11. Boxes mark second quartile, median and third quartile. Whiskers indicate first and fourth quartile. As listed in Table 6 and shown in Figure 11, both the MLP and SOM model on average yield a better prediction than the baseline (ETA-pilot). Only our MLP model shows a smaller spread of the average relative error, compared to the baseline. We conclude that an MLP is better suitable than SOM to estimate the fuel consumption for the given inputs of parameter group D or E in Table 4.

Overall the experiment and development work presented in this phase of “second diamond” shows its feasibility and potential value in using historical big data to predict fuel consumption and support decision making in eco-driving. Although the models, SOMs and MLPs were still not entirely robust, given the data set that we obtained, they are able to outperform the baseline tool used onboard in terms of fuel consumption prediction, as it was essentially using the historical data rather than following the hardcoded programme. By extracting values out of the large amount of data, there is a huge potential to support the development of knowledge of the seafarers and maritime EE optimization operations. At the same time, the impact of using big data analytics may also generate important insights regarding interface design, shore-based management and perhaps future research of ethnography within this domain.

The big data analytics approach, which was initially informed by the ethnographic research in the field, builds up a basis to discuss how we can provide decision support by showing real-time fuel consumption prediction. By the time the ethnographic research and the experiments were done, the research team presented the findings to the shipping company and obtained their positive feedback. They were very interested in the interdisciplinary study. In the initial discussion, they raised some sociotechnical questions that point to the needs for further research, for example, how to make the predictions more robust, what data were further needed and considerations regarding how the management and practitioners shall better accommodate the introduction of big data analytics

onboard ships to enable improved organizational EE practices. This again echoes the complementary nature of thick data and big data as well as showing the value of a pluralistic epistemology in exploratory research.

5. Discussion

5.1. Design of Decision Support Systems and Management Practices

The commercially available EE monitoring toolboxes, like the ETA-pilot systems that are employed on dozens of ferry vessels, are mainly used to display ship sensor information (wind, trim, fuel consumption etc.). A huge problem is that such systems usually tend to assume the display of fuel-consumption data curves equals knowledge creation. However, the ethnographic findings reject this simple assumption, but identify the need to involve context in the process of knowledge mobilization [57]. A major problem is that the design of this existing tool did not really afford the meanings behind the numbers, thus the crews would feel difficult to perform efficient monitoring or flexibly adapt themselves under new circumstances. For example, 20 kg fuel consumption shown per nautical mile is just a number about the rate of burning fuel, but the crews have little understanding about its meaning in EE operations. The information about the ongoing trend to spend more than 5 kg fuel per nautical mile than the “historical average consumption performance” provides probably more support to their decision making. In other words, a better decision support interface design should provide the meanings behind the visualizations to enable direct perception and manipulations, emphasizing the context in which the knowledge could be applied and information could be operationalized.

Confronted with the energy-saving challenges in the design and use of technology identified in the ethnographic study, the research team took the predictive-analytical approaches with big data analytics to explore a path forward. The big data analytics conducted and their evolution in this paper have clearly suggested that, given an operational and environmental profile, machine learning could be feasible to (1) predict fuel consumption in real time and (2) represent the most “similar” journeys and their fuel consumption data (as benchmarks). This is essentially providing a context by attaching meanings to visual data and making it much easier for human operators to understand, e.g., how this maneuvering would affect fuel consumption tendency, or how the energy consumption curve for the same ferry ship under the similar conditions would look like.

One concrete example is given here to show how big data analytics could be employed to support decision making under different weather conditions. Bad weather could lead to more fuel consumption while good weather could help to save fuel. The expected values of fuel consumption under both situations can be considered as useful indicators or measures, to suggest the performance of the real-time fuel consumption, so that the navigators could have an easily understandable metrics to evaluate their own EE practice and possibilities to develop their knowledge and awareness in real-time (or retrospectively).

Figure 12 shows a simplified heuristic design example. The interface plots three curves regarding fuel consumption from time 14:00 to 23:00. Based on the machine-learning algorithm, trips under similar weather conditions were classified from the historical big data set. Then the average fuel consumption values (kg per nautical mile) for “best” and “worst” circumstances are plotted as benchmarks. The dynamic yellow curve describes the real-time fuel consumption, with the dotted line predicting how the consumption is going to be if maintaining the same speed and course under the current weather conditions. In principle the yellow consumption curve should be located between the two other lines (normal performance), but it might go higher than the worst case or lower than the best case, which suggests the “moments of learning” for seafarers and the management.

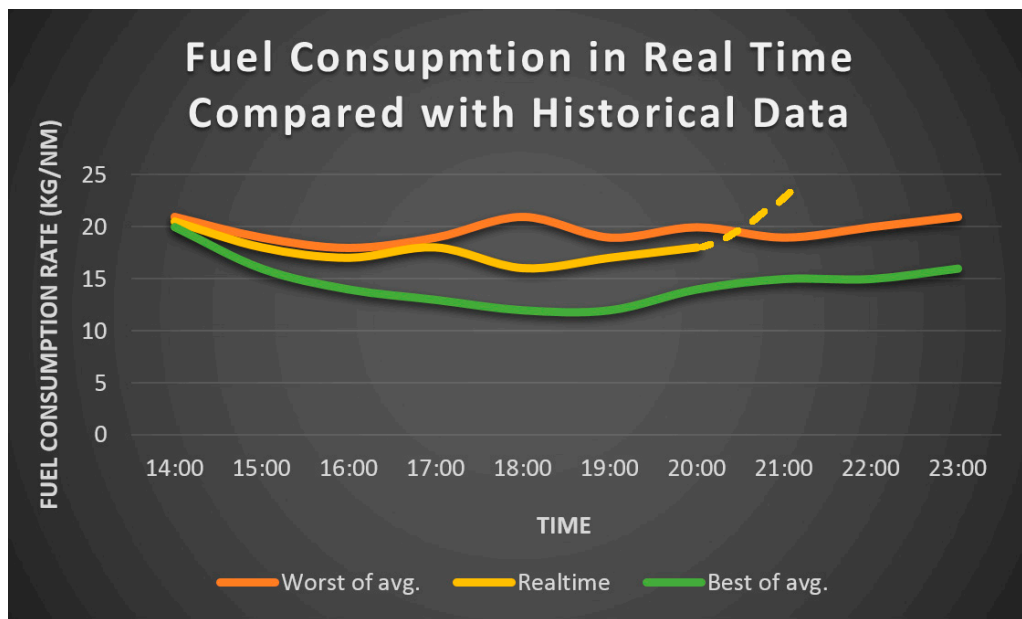


Figure 12. Fuel consumption in real time compared with historical consumption curves.

The idea is not to present a fully-fledged design solution but to suggest a promising direction for system design: by exploiting the power of big data analytics we may directly relieve some of the existing pains and create desired gains. Informed by the ethnographic research, the big data analytics can literally connect the past with the present so that a new sort of knowledge to enable operationalization could emerge (e.g., how a navigator would try to adjust speed or course when he/she was able to see the predictable rise in the curve, or how the dynamic system feedback would allow him/her to form a new understanding about eco-driving and develop capabilities to adapt to the changing environment). This in essence forms a way to address the knowledge gap and influence the quality of human experience, as the users would be provided with “affordance” to be acted upon, an ecological construct that is tightly associated with perception and action [58].

Machine learning with big data analytics cannot only contribute to perception and action at the sharp end, but also might change the fundamental ways of how organizations make decisions and manage EE [59]. The evaluations have revealed that proposed models generalize well on different training data sets from different ships, indicating machine learning’s high-potential role in creating institutionalized knowledge in an organizational context. Addressing the knowledge gaps in organizational decision making is crucial [11], especially for establishing new energy-related policies. In other words, the management needs to see the big picture about what is really going on for the fleet during the voyage and the data shall provide a basis to ground the ship–shore communication and energy related decision making. Introducing the big data analytics to the management practices is an endeavor to create a holistic understanding and establish a bottom-up approach for energy management (e.g., why certain groups of crew members on average consume more fuel than the other groups, what challenges they may encounter, what is needed to be done to facilitate their eco-driving, what measures for EE shall be considered and adopted, etc.). The aim is to take advantage of the power of big data and enable organizational knowledge mobilization and development, which is a vision that comes along with our experiments using big data analytics.

5.2. Limitations and Future Research

Although the presented machine-learning models have good performance with respect to the ETA-pilot, the accuracies of the presented machine-learning models were not ideal and future research is needed to test more machine-learning algorithms. The trained SOM model yields the worst robustness,

compared to the MLP and to the existing tools, probably because SOMs require more data samples to converge into a state that yields a similar performance to the regression models. The collection of roughly 60 ship-months should be extended to cover a longer time period per ship. Still the evaluation showed that both MLP and SOM generalize well on different given data sets. This is largely because the appropriate hyper parameters enabled the models to be generalizable to some extent when trained on different ferries, meaning that the models' predictive performance is not bound to one ship. However, it should also note that the current hyperparameters might not work that well for ships that do not sail on the same routes, such as tankers and dry bulkers instead of ferries. Therefore, there is a huge unexploited potential to be trained on different ships and routes in the future.

Hull fouling, maintenance costs could also be included in the model to predict real voyage costs (which is a composite of more factors than only fuel costs). Current data do not include critical data points such as trim or draught. The specific type of fuel that was used is also unknown but might have a big impact on the engines' performance. Ship sensors work at a higher time resolution than 10 min, but only a 10-min average is stored on hard disks. An increase of the time resolution of stored sensor data is a promising way to collect more data in order to train better models. For a holistic energy-efficiency analysis, one has to include long-term costs as for example maintenance costs, that also depend on the operational profile of the engines. The slow-steaming in shipping might save fuel but can increase maintenance overhead in cases in which the engines were designed for a different speed. All these create a vast space for big data analytics.

On the other hand, the work in big data analytics and its derived discussions above also indicate there is much space for future ethnographic research, which can be employed to discover the meanings behind big data analytics [28]. The words of the sociologist William Bruce Cameron is that not everything that counts can be counted, and not everything that can be counted counts. If human users can gain knowledge (as a social construct) during the use of the technology, would the technology support the crews' social interactions (e.g., exchanging ideas to save more fuel)? How would these types of data-driven predictive tools influence communication in the communities of practice, or the communication between ship and shore, considering collaborative learning is critical for maritime EE optimization practices [24]? More importantly, how is the known issue of lack of trust between ship and shore organization [26] likely to be influenced by the introduction of the human-AI decision systems also remains to be explored, as the intelligent system will be more involved in the ship operations but still be put under "administration" of the shore organization? It is important to consider the role of advanced technology and avoid the erosion of the trust and communication link between ship and shore. Technology should support the whole organization to evaluate performance and understand what concrete operations could constitute "best practice" or "worst practice" of EE optimization. A feasible research direction is to understand how the big data analytics could shape the navigators' work in the actual field and how would the management practices need to accommodate this. For designers, this is also a question pertinent to "work as imagine" vs. "work as done" [60]. These questions are something the big data techniques fall short in answering and we need ethnography to deliver a deeper understanding of the organizational context in which big data technology gets utilized, so that better organizational practice (involving big data based systems) could be proposed and validated later on.

In this paper, ethnographic insights emerge as the departure point for big data analytics and the big data analytics indicates more possibility for ethnography to understand the context researched. The "iteration" or "combination" of the ethnography and big data analytics may illustrate the complementary natures of these two methodologies [44] and the need to develop a richer methodological framework for future maritime research [61].

6. Conclusions

This interdisciplinary study provides a detailed description of the use of both the ethnographic research method and big data analytics to explore the phenomena under study and proposes a way forward. It uses a fuel monitoring system used onboard and maritime EE as the applied context. The ethnographic study sets the departure point of the research on big data analytics and the follow-up discussion. It identifies that the tools available are not suitable for the task and are not fully understood by the crew. A lack of a performance evaluation and subsequent analytical approaches have created significant operational barriers in reducing fuel consumption, whereas the observed shipping company also lacked the competence to deal with the energy-related data. This informs the need for using big data analytics that have the computational power to analyze data and support the learning and understanding of actual fuel consumption without adding extra operational workloads upon the ship operations personnel. Furthermore, the ethnographic research also suggests some strategic direction that is worth further investigation as well as some features that may be playing significant roles in machine-learning modelling, such as speed, course, locations, wind, fuel consumption. Therefore, the introduction of big data analytics is situated in a particular context (rather than “decontextualization” typical of most technical studies or controlled environments). This paper did not stop after the ethnographic study but continues to take further steps to explore feasible solutions. A detailed process was provided regarding how the big data analytics were applied to develop predictive-analytical algorithms by using real ship data and evaluating supervised and unsupervised machine-learning techniques. The evaluation showed that based on the training data set the models output more accurate predictions on fuel consumption on average, compared with the existing tool used onboard ships, suggesting a huge potential for system design that can support decision making and improve EE optimization practice. The work around AI also raises more questions that direct future ethnographic research.

The purpose to employ the ethnographic study and big data analytics in this research project is not to produce something that can immediately replace the technology used onboard but to demonstrate the process of an interdisciplinary research approach in which big data analytics could be initially inspired and informed by ethnographic research. That is the value of the pluralistic epistemology [62] as multiple complementary perspectives (big data and thick data) are taken to explore the connections between disciplinary knowledge. This completes the closure of a system design in the double diamond framework and forms a body of holistic knowledge of interdisciplinary applied research in the maritime EE domain.

As discussed in the paper, more thick data might still be required to reveal the social context and generate meaningful insights behind big data, which would form an iterative research paradigm. Our interdisciplinary study has shed light on what should be considered when designing decision support systems as well as management practices from the perspectives of human–computer interaction and ship–shore interaction. The complementary nature of big data and ethnography [44] revealed might suggest a methodologically reflexive path forward for future maritime energy research.

Author Contributions: Conceptualization, Y.M.; methodology, Y.M. and T.S.; software, T.S.; validation, Y.M., T.S., M.L. and S.N.M.; formal analysis, Y.M. and T.S.; investigation, Y.M. and T.S.; resources, T.S. and M.L.; data curation, T.S.; writing—original draft preparation, Y.M.; writing—review and editing, Y.M., T.S., M.L. and S.N.M.; visualization, Y.M. and T.S.; supervision, M.L. and S.N.M.; project administration, M.L. and S.N.M. All authors have read and agreed to the published version of the manuscript.

Funding: The work is a part of Y.M.’s Ph.D. research program and T.S.’s master thesis work (internal circulation of this work) from Chalmers University of Technology and the Karlsruhe Institute of Technology, thus this research received no external funding.

Acknowledgments: The authors want to thankfully acknowledge the shipping company for support to this work.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Measurements of on-board ship sensors.

Data Point Name	Explanation	Comment
date-time	local time	
duration	unknown	not used
fuel mass flow HFO ME	heavy fuel oil, main engines	group 1 and 2
fuel mass flow MDO ME	marine diesel oil, main engines	group 1 and 2
fuel mass flow HFO AUX	heavy fuel oil, auxiliary engines	group 1 and 2
fuel mass flow MDO AUX	marine diesel oil, auxiliary engines	group 1 and 2
fuel mass flow per nm fuel mass flow total (kg/h)	tons per nautical mile	
AWA	apparent wind angle in degrees	relative to ship's heading
AWS knots	apparent wind speed	
AWS m/s	apparent wind speed	
COG	course over ground in degrees	magnetic/GPS
CDIR	current direction	magnetic/GPS
CSPD knots	current speed	
CSPD (m/s)	current speed	
displacement (mt)	amount of displaced water	not used
DBT (m)	water depth	not used
distance (nm)	usage unknown	ranges between 0 and 2
draft aft (m)	depth of ship, measured at the stern of the vessel	no used
draft fore (m)	depth of ship, measured at the bow of the vessel	no used
HTD (rad)	heading true	
heel (deg)	roll, heeling angle (upright = 0)	
LWS(knots)	longitudinal water speed	
LWS froude (knots)	froude speed through water	not used
PowerConsBT (kW)	power consumption for bow thruster	separate for 4 thrusters
PowerProdAux (kW)	power production aux engines	separate for 5 engines
PowerProdShore (kW)	power intake shore connection	
PowerProdShaftGen (kW)	power production shaft generator	separated for each ME
propulsion pitch (%)	% of maximum propeller pitch	separated for each propeller
propulsion RPM	propulsion of an engine	separated for each shaft
propulsion power (%)	% of maximum power output	separated for each shaft
SFOC MDO (g/kWh)	specific oil consumption	no data
SFOC HFO (g/kWh)	specific oil consumption	no data
SOG (knots)	speed over ground	
SOG Froude (knot)	froude speed over ground	
SWT (knot)	speed through water	
sea water temperature (°C)		
TWDIR (degrees)	true wind direction	
TWS (degrees)	true wind speed	not used
TWSPD (degrees)	true wind speed	
trim (m)	difference of draфтаft/-fore	not used
trim dynamic (m)	difference of draфтаft/-fore	not used

References

- UNCTAD. *Review of Maritime Transport 2019*; United Nations: New York, NY, USA, 2019.
- IMO. *Initial IMO Strategy on Reduction of GHG emissions from Ships, in Resolution MEPC*; International Maritime Organisation: London, UK, 2018; Volume 304.
- Johnson, H.; Styhre, L. Increased energy efficiency in short sea shipping through decreased time in port. *Transp. Res. Part A Policy Pract.* **2015**, *71*, 167–178. [[CrossRef](#)]

4. IMO. *Third IMO Greenhouse Gas Study: Executive Summary and Final Report*; International Maritime Organization (IMO): London, UK, 2014.
5. Bännstrand, M. *Study on the Optimization of Energy Consumption as Part of Implementation of a Ship Energy Efficiency Management Plan (SEEMP)*; International Maritime Organisation (IMO): London, UK, 2016.
6. ABS. *Ship Energy Efficiency Measures Advisory*; ABS: Houston, TX, USA, 2013.
7. Winkel, R.; Bos, A.v.d.; Weddige, U. *Study on Energy Efficiency Technologies for Ships: Inventory and Technology Transfer (Final Report)*; Publications Office of the European Union: Luxembourg, 2015.
8. Viktorelius, M.; Lundh, M. Energy efficiency at sea: An activity theoretical perspective on operational energy efficiency in maritime transport. *Energy Res. Soc. Sci.* **2019**, *52*, 1–9. [[CrossRef](#)]
9. Johnson, H. *Search of Maritime Energy Management*; Chalmers University of Technology: Göteborg, Sweden, 2016.
10. Banks, C.; Turan, O.; Incecik, A.; Lazakis, I.; Lu, R. Seafarers’s current awareness, knowledge, motivation and ideas towards low carbon-energy efficient operations. *J. Shipp. Ocean Eng.* **2014**, *2*, 11–20.
11. Kitada, M.; Ölçer, A. A Contemporary Outlook on Human Element in Energy Efficient Seaborne Transportation. In Proceedings of the 16th IAMU Annual General Assembly, Opatija, Croatia, 7–10 October 2015.
12. Viktorelius, M.; Lundh, M. The Role of Distributed Cognition in Ship Energy Optimization. In *Energy Efficient Ships*; The Royal Institution of Naval Architects: London, UK, 2016.
13. Acciaro, M.; Hoffmann, P.N.; Eide, M.S. The Energy Efficiency Gap in Maritime Transport. *J. Shipp. Ocean Eng.* **2013**, *3*, 1–10.
14. Jafarzadeh, S.; Utne, I.B. A framework to bridge the energy efficiency gap in shipping. *Energy* **2014**, *69*, 603–612. [[CrossRef](#)]
15. Poulsen, R.T.; Johnson, H. The logic of business vs. the logic of energy management practice: Understanding the choices and effects of energy consumption monitoring systems in shipping companies. *J. Clean. Prod.* **2016**, *112*, 3785–3797. [[CrossRef](#)]
16. Johnson, H.; Johansson, M.; Andersson, K.; Södahl, B. Will the ship energy efficiency management plan reduce CO2 emissions? A comparison with ISO 50001 and the ISM code. *Marit. Policy Manag.* **2013**, *40*, 177–190. [[CrossRef](#)]
17. Johnson, H.; Andersson, K. Barriers to energy efficiency in shipping. *WMU J. Marit. Aff.* **2016**, *15*, 79–96. [[CrossRef](#)]
18. Johnson, H.; Johansson, M.; Andersson, K. Barriers to improving energy efficiency in short sea shipping: An action research case study. *J. Clean. Prod.* **2014**, *66*, 317–327. [[CrossRef](#)]
19. Lützen, M.; Mikkelsen, L.L.; Jensen, S.; Rasmussen, H.B. Energy efficiency of working vessels—A framework. *J. Clean. Prod.* **2017**, *143*, 90–99. [[CrossRef](#)]
20. Poulsen, R.T.; Sampson, H. ‘Swinging on the anchor’: The difficulties in achieving greenhouse gas abatement in shipping via virtual arrival. *Transp. Res. Part D Transp. Environ.* **2019**, *73*, 230–244. [[CrossRef](#)]
21. Rasmussen, H.B.; Lützen, M.; Jensen, S. Energy efficiency at sea: Knowledge, communication, and situational awareness at offshore oil supply and wind turbine vessels. *Energy Res. Soc. Sci.* **2018**, *44*, 50–60. [[CrossRef](#)]
22. Hansen, E.K.; Rasmussen, H.B.; Lützen, M. Making shipping more carbon-friendly? Exploring ship energy efficiency management plans in legislation and practice. *Energy Res. Soc. Sci.* **2020**, *65*, 101459. [[CrossRef](#)]
23. Kataria, A.; Holder, E.; Praetorius, G.; Baldauf, M.; Schröder-Hinrichs, J.U. Exploring Bridge-Engine Control Room Collaborative Team Communication. *Int. J. Mar. Navig. Saf. Sea Transp.* **2015**, *9*, 169–176. [[CrossRef](#)]
24. Man, Y.; Lundh, M.; MacKinnon, S.N. Maritime Energy Efficiency in a Sociotechnical System: A Collaborative Learning Synergy via Mediating Technologies. *Int. J. Mar. Navig. Saf. Sea Transp.* **2018**, *12*, 239–250. [[CrossRef](#)]
25. Viktorelius, M. Adoption and use of energy-monitoring technology in ship officers’ communities of practice. *Cognit. Technol. Work* **2019**, 1–13. [[CrossRef](#)]
26. Sampson, H.; Turgo, N.; Acejo, I.; Ellis, N.; Tang, L. ‘Between a Rock and a Hard Place’: The Implications of Lost Autonomy and Trust for Professionals at Sea. *Work Empoly. Soc.* **2019**, *33*, 648–665. [[CrossRef](#)]
27. Latzko-Toth, G.; Bonneau, C.; Millette, M. Small Data, Thick Data: Thickening Strategies for Trace-Based Social Media Research. In *The SAGE Handbook of Social Media Research Methods*; SAGE Publishing: Thousand Oaks, CA, USA, 2017; pp. 199–214.

28. Wang, T. Why Big Data Needs Thick Data. 2016. Available online: <https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7> (accessed on 27 March 2019).
29. Dourish, P. Implications for design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, Montréal, QC, Canada, 22–27 April 2006; pp. 541–550.
30. Norman, D. The Research-Practice Gap. *Interactions* **2010**, *17*, 9–12. [[CrossRef](#)]
31. Grech, M.R. *Human Factors in the Maritime Domain*/Michelle Rita Grech; Horberry, T., Koester, T., Eds.; CRC Press: Boca Raton, FL, USA, 2008.
32. Yan, X.; Wang, K.; Yuan, Y.; Jiang, X.; Negenborn, R.R. Energy-efficient shipping: An application of big data analysis for optimizing engine speed of inland ships considering multiple environmental factors. *Ocean Eng.* **2018**, *169*, 457–468. [[CrossRef](#)]
33. Perera, L.P.; Mo, B. Machine Intelligence Based Data Handling Framework for Ship Energy Efficiency. *IEEE Trans. Veh. Technol.* **2017**, *66*, 8659–8666. [[CrossRef](#)]
34. Leifsson, L.; Sævarsdóttir, H.; Sigurðsson, S.P.; Vésteinsson, A. Grey-box modeling of an ocean vessel for operational optimization. *Simul. Model. Pract. Theory* **2008**, *16*, 923–932. [[CrossRef](#)]
35. Ahlgren, F. Reducing ships' fuel consumption and emissions by learning from data. In *Kalmar Maritime Academy*; Linnaeus University Press: Kalmar, Sweden, 2018.
36. Coraddu, A.; Oneto, L.; Baldi, F.; Anguita, D. Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. *Ocean Eng.* **2017**, *130*, 351–370. [[CrossRef](#)]
37. Victor, P.d.C.; Joakim, D.; Ana, N. What is maritime navigation? Unfolding the complexity of a Sociotechnical System. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, New Orleans, LA, USA, 5–8 March 2017; Volume 61, pp. 267–271.
38. Man, Y.; Lundh, M.; MacKinnon, S.N. Facing the New Technology Landscape in the Maritime Domain: Knowledge Mobilisation, Networks and Management in Human-Machine Collaboration. In Proceedings of the 9th International Conference on Applied Human Factors and Ergonomics, Washington, DC, USA, 24–28 July 2019; Springer International Publishing: Orlando, FL, USA, 2019.
39. Wang, H.; Osen, O.L.; Li, G.; Li, W.; Dai, H.N.; Zeng, W. Big data and industrial Internet of Things for the maritime industry in Northwestern Norway. In Proceedings of the TENCON 2015-2015 IEEE Region 10 Conference, Macao, China, 1–4 November 2015.
40. Rødseth, Ø.J.; Perera, L.P.; Mo, B. Big Data in Shipping—Challenges and Opportunities. In Proceedings of the 15th International Conference on Computer and IT Applications in the Maritime Industries (COMPIT'16), Lecce, Italy, 9–11 May 2016.
41. Mayer-Schönberger, V.; Cukier, K. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*; John Murray: London, UK, 2013.
42. Viktorelius, M. The human and social dimension of energy efficient ship operation. In Proceedings of the International Conference on Maritime Energy Management, Malmö, Sweden, 24–25 January 2017.
43. Patton, M.Q. *Qualitative Research & Evaluation Methods*; SAGE: Thousand Oaks, CA, USA, 2002; p. 598.
44. Charles, V.; Gherman, T. Big Data Analytics and Ethnography: Together for the Greater Good. In *Big Data for the Greater Good*; Emrouznejad, A., Charles, V., Eds.; Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 19–33.
45. Man, Y.; Lützhöft, M.; Costa, N.A.; Lundh, M.; MacKinnon, S.N. Gaps Between Users and Designers: A Usability Study About a Tablet-Based Application Used on Ship Bridges. In Proceedings of the AHFE 2017 International Conference on Human Factors in Transportation, Los Angeles, CA, USA, 17–21 July 2017; Stanton, N.A., Ed.; Springer International Publishing: Berlin/Heidelberg, Germany, 2018; pp. 213–224.
46. Banathy, B.H. *Designing Social Systems in a Changing World*; Springer: Berlin/Heidelberg, Germany, 1996.
47. Council, B.D. What Is the Framework for Innovation? Design Council's Evolved Double Diamond. 2004. Available online: <https://www.designcouncil.org.uk/news-opinion/what-framework-innovation-design-councils-evolved-double-diamond> (accessed on 5 March 2020).
48. Brewer, J. *Ethnography*; McGraw-Hill Education: London, UK, 2000.
49. Schensul, S.L.; Schensul, J.J.; LeCompte, M.D. *Essential Ethnographic Methods: Observations, Interviews, and Questionnaires*; AltaMira Press: Lanham, MD, USA, 1999; Volume 2.
50. Anan, T.; Higuchi, H.; Hamada, N. New Artificial Intelligence Technology Improving Fuel Efficiency and Reducing CO2 Emissions of Ships through Use of Operational Big Data. *Fujitsu Sci. Tech. J.* **2017**, *53*, 23–28.

51. Senin, P. Z-Normalization of Time Series. 2016. Available online: https://jmotif.github.io/sax-vsm_site/morea/algorithm/znorm.html (accessed on 21 March 2019).
52. Rosenblatt, F. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*; Spartan Books: Washington, DC, USA, 1962.
53. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Cogn. Modeling* **1988**, *5*, 1. [[CrossRef](#)]
54. Kingma, D.P.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980. Available online: <https://arxiv.org/abs/1412.6980v9> (accessed on 21 March 2019).
55. Kiviluoto, K. Topology preservation in self-organizing maps. In Proceedings of the International Conference on Neural Networks (ICNN'96), Washington, DC, USA, 3–6 June 1996.
56. Martin, J.; Meredith, M.A.; Ahmad, K. Modeling multisensory enhancement with self-organizing maps. *Front. Comput. Neurosci.* **2009**, *3*, 8. [[CrossRef](#)]
57. Parent, R.; Roy, M.; St-Jacques, D. A systems-based dynamic knowledge transfer capacity model. *J. Knowl. Manag.* **2007**, *11*, 81–93. [[CrossRef](#)]
58. Flach, J.M.; Tanabe, F.; Monta, K.; Vicente, K.J.; Rasmussen, J. An Ecological Approach to Interface Design. In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago, IL, USA, 5–9 October 1998.
59. PwC. The human factor: Working with machines to make big decisions. In *PwC's Global Data and Analytics 2016: Big Data*; PwC: New York, NY, USA, 2016.
60. Hollnagel, E. Why is work-as-imagined different from work-as-done. In *Resilient Health Care: The Resilience of Everyday Clinical Work*; Wears, R.L., Hollnagel, E., Braithwaite, J., Eds.; Ashgate: Farnham, UK, 2015; Volume 2.
61. Elish, M.C.; Boyd, D. Situating Methods in the Magic of Big Data and Artificial Intelligence. *Commun. Monogr.* **2017**, *85*, 57–80. [[CrossRef](#)]
62. Man, Y.; Lundh, M.; MacKinnon, S.N. Towards a Pluralistic Epistemology: Understanding Human-Technology Interactions in Shipping from Psychological, Sociological and Ecological Perspectives. *Int. J. Mar. Navig. Saf. Sea Transp.* **2018**, *12*, 795–811. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).