Marine Engine Centered Data Analytics for Ship Performance Monitoring

Perera, Lokukaluge P.[1]
Norwegian Marine Technology Research Institute (MARINTEK)
Trondheim, Norway
Prasad.Perera@marintek.sintef.no

Mo, Brage
Norwegian Marine Technology Research Institute (MARINTEK)
Trondheim, Norway
Brage.Mo@marintek.sintef.no

ABSTRACT

This study proposes marine engine centered data analytics as a part of the ship energy efficiency management plan (SEEMP). The SEEMP enforces various emission control measures to improve ship energy efficiency by considering vessel performance and navigation data. The proposed data analytics is developed in the engine-propeller combinator diagram (i.e. one propeller shaft with a direct drive main engine). Three operating regions from the initial data analysis are under the combinator diagram noted to capture the shape of these regions by the proposed data analytics. The data analytics consists of implementing Gaussian Mixture Models (GMMs) to classify the most frequent operating regions of the main engine. Furthermore, the Expectation Maximization (EM) algorithm calculates the parameters of GMMs. This approach also named as a data clustering algorithm facilitates an iterative process for capturing the operating regions of the main engine (i.e. in the combinatory diagram) with the respective mean and covariance matrices. Hence, these data analytics can monitor ship performance and navigation conditions with respect to engine operating regions as a part of the SEEMP. Furthermore, development of advanced mathematical models for ship performance monitoring within the operational regions (i.e. data clusters) of marine engines is expected.

[1] Corresponding author.

INTRODUCTION

Energy Efficiency Management Plan

Modern vessels are equipped with various onboard sensors and data acquisition (DAQ) systems to collect ship performance and navigation information. This information is data sets collected and analyzed to evaluate ship performance under various sea going conditions. These data analysis methods, so called "data analytics", and respective results should be a part of the ship energy efficiency management plan (SEEMP). The proposed analytics can develop under the SEEMP to draw conclusions on ship energy efficiency from the respective data sets. Such data analytics can play an important role in commercial ship operations in the future years as a part of the SEEMP ([1]-[2]). The SEEMP as a mandatory mechanism in shipping enforces vessels to improve operational conditions and implement technology advancements for more energy efficient shipping fleets. The energy efficiency operational indicator (EEOI) [3] is a benchmark level, a vessel performance monitoring mechanism, for the SEEMP. Therefore, vessel performance and navigation information under the proposed data analytics supports to achieve the objectives in ship energy efficiency assigned by the SEEMP.

The SEEMP has following phases: planning, implementation, monitoring, self-evaluation and improvements. The planning phase is the most important step in the SEEMP. Goal setting should be with the human resource development strategies of the shipping companies coordinated at this phase. The implementation phase is an important step in the SEEMP that uses appropriate established procedures. The monitoring phase consists of developing the required condition monitoring (CM) facilities (i.e. sensors, DAQ systems, data storage and communication) [4, 5]. The final step of the SEEMP is the self-evaluation phase conducting voluntary reporting and review processes. Furthermore, the documentation of lessons learned for further improvements should also be part of this phase.

A goal for the SEEMP is to control medium term exhaust emissions by improving ship energy efficiency. To achieve this the ship crew should have proper knowledge and training on energy management approaches. Such energy management approaches improve vessel performance that various data analytics should support. The proposed data analytics based on ship performance and navigation data increases the understanding of efficient ship operating conditions. This study proposes to use marine engine centered data analytics in the SEEMP to observe vessel performance. Furthermore, the proposed data analytics at the ship operation phase can identify optimal engine-propulsion conditions of vessels [6].

Appropriate engine-propulsion configurations (i.e. optimal operating conditions) in ships can reduce power/fuel consumption and exhaust emissions, significantly. This study proposes to use the combinator diagram (i.e. the relationship between main engine (ME) power and shaft speed) to identify appropriate engine-propulsion configurations, where the respective data analytics is implemented. This can also identify the most frequent operating regions (i.e. speed and power conditions) of marine engines. A data set of vessel performance and navigation information of a selected ship (i.e. one propeller shaft with a direct drive main engine) is used to develop these data analytics. An overview of the engine-propulsion combinator diagram is in the following section discussed.

Engine Propeller Combination

Optimal engine-propulsion configurations are at the design phase of vessels selected to meet ship operational and navigation requirements. However, such configurations can be at the ship operation phase challenged due to various environmental factors. Engine-propeller interactions should be in such navigation situations monitored and analyzed to evaluate ship performance. The engine-propulsion interactions are often studied using the combinator diagram. A general overview of the engine-propulsion combinator diagram is in Figure 1 presented and consists of a relationship between main engine power (in a log scale) and shaft/propeller speeds. This is from consideration of various engine-propeller combinator diagrams in marine engines derived. A direct drive situation, the main engine (ME) directly connects to a propeller shaft that drives the propeller (i.e. fixed-pitch-propeller). In general, such diagram represents engine-propeller operational regions with respect to various ship navigation situations [7]. Therefore, modern integrated bridge systems are also equipped with such onboard combinator diagrams to evaluate vessel performance in various environmental conditions. Furthermore, the respective engine fuel consumption (i.e. specific fuel consumption (SFC)) can be into these combinator diagrams incorporated to monitor vessel performance and emissions, continuously.

The engine and propeller operating regions in a combinator diagram are limited by various ship navigation constrains and some selected features are presented in

Figure 1. The closed area (i.e. marked by green color lines) represents the possible engine-propeller region with respective various vessel operating constrains (i.e. engine speed, power and operating limits). However, the engine-propeller operating regions also depend on engine manufacturer specifications, therefore various combinator diagrams are in the literature presented. The maximum engine power limit (C2) (that can be 100% of respective engine power) of the engine is presented in the same figure. One should note that the specified maximum continuous rating (MCR) point (P1) of the engine is located at the intersection of the 100 % of engine power and speed axes. Furthermore, this engine-propeller operating region can be further into several circular regions (D1) divided of specific fuel consumption (SFC). The smallest SFC region (i.e. the smallest radius) represents the optimal fuel consumption rate of the engine. Therefore, the propeller operating points in vessels should be close to the optimal SFC region to reduce the fuel consumption and improve vessel performance.

(E1) represents a light running fixed-pitch-propeller (FPP) curve under clean hull and propeller conditions in calm water. The propeller design point (P2) and alternative propeller design point (P3) are at the same line located. One should note that (P2) can move towards (P3) due to the sea margin (F1). However, (P2) can shift towards the heavy FPP curve (E2) due to fouled hull and propeller conditions and/or rough weather situations. Therefore, a light running engine, (P2), can move towards a heavy running engine, (P4), due to such conditions. (P4) can shift toward (P5) due to the sea margin (F1) under heavy running conditions in the main engine. Furthermore, (P5) can move towards the specified MCR point for propulsion, (P6), due to the additional engine margin (F2) (up to 20% of the engine power). If the main engine has a shaft generator (i.e. power can be taken out of the system for electricity production), then (P6) can move towards (P1) under PTO (i.e. power take out) situations. Even though the ultimate heavy FPP curve (E3) can intercept (P1) as presented in the figure, the location of (P1) can vary due to main engine specifications. If the main engine has a shaft motor, then additional power can be injected into the propulsion system (i.e. power take in (PTI)), where (P6) can move away from (P1). Therefore, engine-propeller combinator diagrams may have additional variations due to these reasons.

The recommended optimal operating point (P7) represents the lowest SFC value (i.e. the optimal fuel consumption) for the respective engine. Therefore, propeller should be near this point operated to improve vessel performance, as mentioned before. However, the engine operating point can move towards (P8), the continuous service rating of the engine, as per vessel operational requirements (i.e. heavy running situations). Therefore, operating points of propeller (i.e. FPP) can vary along an approximately straight line with respect to main engine running conditions under similar weather conditions. A considerable ship speed reduction under engine heavy running conditions increases the SFC. One should note that engine operating regions and accepted limits discussed above vary due to requirements from their manufacturers and ship owners. Therefore, additional variations on combinator diagrams is in the shipping industry are as mentioned before noted. However, the selection of optimal engine-propeller operating regions with respect to the most frequent operating regions can play an important role in ship performance monitoring as a part of the SEEMP.

The main objective in this study is to develop data analytics to capture the most frequent main engine operating regions (i.e. main engine power and shaft speed). Similarly, such engine-propeller operating regions can be part of the evaluation of ship performance under various mathematical models. That can be done by monitoring ship performance and navigation information in real-time and evaluating those parameters with respect to main engine operational regions, where the respective SFC should be minimized. In general, marine engines operate around selected RPM values (i.e. operating modes), where each main engine has several mean RPM values to use to achieve different ship speeds required.

These operating modes also relate to engine loading conditions. Hence, each marine engine operates around several mean RPM values to achieve required ship speeds. The engine related operating conditions can by ship performance and navigation data that are collected by onboard sensors and data acquisition systems (DAQs) be captured in such situations. Data sets in ship performance and navigation parameters should be further along the engine operating regions classified to identify such vessel operating modes, experimentally. Hence, the proposed data analytics consist of these data classification and identification steps to visualize main engine modes.

Various data analytics are often possible to use to facilitate decision support tools in the maritime and offshore industries [8]. These decision support tools are on large data sets based and that may introduce additional challenges in the respective data analyses. In general, such large data sets of ship performance and navigation information consist of some erroneous data conditions due to sensors and DAQ noise. However, these erroneous conditions can also be identified, effectively by data classification (i.e. data clustering) approaches, where unusual data patterns (i.e. erroneous conditions) are more visible in smaller data sets. Hence, such erroneous conditions in ship performance and navigation parameters are possible to remove and create cleaner data sets [9]. Hence, the improved data sets (i.e. without erroneous data conditions) are developed by marine engine centered data analytics to evaluate ship performance under the SEEMP [10, 11].

Ship Performance Monitoring

Ship performance and navigation data in a selected vessel are in this study considered and the vessel particulars are presented in Table 1. A data set of three parameters from the above vessel categories: shaft speed, main engine (ME) Power and fuel consumption. The following parameter ranges are in this data analysis: ME Power from 3000 (kW) to 8000 (kW) and Shaft Speed from 80 (RPM) to 120 (RPM). This is the most frequent engine-propeller operating region identified in the combinator diagram, initially. The data points around the zero speed-power values that represent slow moving conditions (i.e. maneuvering/berthing conditions) of the vessel are removed from the data analysis, therefore considerable sensor and DAQ noise conditions (i.e. erroneous data intervals) are filtered by this step. Approximately 50% of the data points are in this data set removed because that relate to maneuvering/berthing conditions of the vessel.

An initial statistical analysis of the selected parameters that relates to general engine-propeller operating situations as presented in Figure 2. The top, middle and bottom plots of the figure are histograms of shaft speed, ME power, and fuel consumption, respectively. Three frequent operating regions are in this statistical analysis observed. That shows the main engine operates around three mean RPM regions that create three mean ME power and fuel consumption regions. As the next step, these two parameters (i.e. ME power and shaft speed) are combined to develop an engine-propeller combinator diagram in a high dimensional space and the results are presented in Figure 3. The bottom plot of the same figure represents a histogram of the ME power and shaft speed combined values. One should note that the same engine operating regions (see Figure 2) are in this combinator diagram also noted. The top-left plot of the same figure represents the contours of the previous plot. The top-right plot of the same figure represents the same contours with the respective fuel computation values. In general, high ME power-shaft speed regions consist of high fuel consumption values and vice versa. However, high fuel consumption values in low ME power-shaft speed regions can be in some situations due to rough ocean conditions also noted.

The same contour plot with respect to relative (Rel.) wind speed is in the top left plot of Figure 4 presented. An assumption is that the relative wind conditions relate to encountered sea states in the respective ship route [4]. The same results show that the shaft speed of the ME reduces significantly for the same engine power because of high wind speeds (i.e. higher engine loading conditions). Furthermore, the propeller is rotating at relatively slow speeds in such situations, where ship speed is also degraded.

The same contour plot with respect to speed through water (STW) of the vessel is in the top right plot of Figure 4 presented. These two plots show that ME power increments can improve STW and rough weather conditions (i.e. high wind speeds) can degrade the same STW, significantly. The same contour plots with respect to vessel trim and average (avg.) draft values are presented in the bottom left and right plots of Figure 4, respectively. One should note that these avg. draft values relate to the loading conditions of the vessel, therefore appropriate trim values reduce ship resistance during vessel navigation.

Another view of the engine-propeller combinator diagram is in Figure 5 presented. The respective operating patterns of the propeller are as straight lines identified in this figure and those are as gray lines marked. These straight lines are during a real-time simulation of the same data set observed. The ME power axis (i.e. y-axis) is presented in a log scale to improve the visibility of the respective engine-propeller data. The vertical gray lines represent various ME power values with constant shaft speed situations. However, the inclined gray lines represent continuous ME operating situations in each voyage segment under varying sea conditions. One should note that these inclined lines are approximately similar to the fixed-pitch-propeller (FPP) lines in Figure 1 (i.e. E1, E2, and E3). These results confirm that the FPP operating points can vary along approximately straight lines with respect to ME running conditions. The same combinator diagram with the vessel STW values is in Figure 6 presented. The results show that vessel STW is degrading along the FPP lines (i.e. for the same engine power, the shaft speed is degrading) from right to left. Therefore, an overview of various interactions among main engine, propeller, ship resistance, and environmental conditions are in such combinator diagrams observed. As the final step in this study, the three operating regions observed in Figure 3 are identified by the proposed data analytics.

DATA ANALYTICS

The proposed marine engine centered data analytics is under the engine-propeller combinator diagram derived. That consists of implementing Gaussian Mixture Models (GMMs) with the Expectation-Maximization (EM) algorithm for clustering the respective data points that relate to the engine operating regions.

Gaussian Mixture Models

Gaussian Mixture Models (GMMs) are extensively for transportation systems used to develop multiple statistical models with large data sets [12]. This study proposes to use the same approach to identify the most frequent operating regions of the engine-propeller combinator diagram. The respective advantages by using GMMs with the EM algorithm are in this section discussed. In general, marine engines operate around several mean RPM values that are often as engine modes categorized. GMMs with the EM algorithm can capture those modes effectively, since the respective engine operation data consist of Gaussian type distributions. One should note that such data sets might distribute in a high dimensional space with respect to other ship performance and navigation parameters. Furthermore, GMMs with the EM algorithm identify not only mean values but also covariance values (i.e. the shape of the engine-operating region) of such engine operating modes (i.e. data distribution) with less supervision.

When the same data set should also under additional ship performance and navigation parameters be classified, the same GMMs can be expanded for a high dimensional space, as required.

The respective parameters of the GMMs (i.e. data clustering) are calculated by an iterative process (i.e. EM algorithm). Each frequent operating region of the engine-propeller combinator diagram is by a GMM represented and the respective parameters are by the EM algorithm calculated. Therefore, each GMM of the combinator diagram has its own mean and covariance values that are at the initial step approximated. Each GMM (i.e. each data set) calculates the respective mean and covariance values under the EM algorithm during the data clustering process. Several independent multivariate Gaussian distributions (i.e. GMM) in the engine-propeller combinatory diagram are during this process introduced.

Level-Two Heads

The respective data points for each GMM are by the EM algorithm [13] assigned, which consists of two iterative levels: expectation and maximization. The EM algorithm is as an effective iterative procedure used for maximum likelihood estimation (MLE). Therefore, it is also to calculate the respective model parameters of the GMMs, as mentioned before used. In the expectation step (i.e. E-step), the probability that each data point belongs to the respective data cluster (i.e. GMM) is evaluated. In the maximization step (i.e. M-step), that data point is in the respective data cluster (i.e. GMM) accommodated with respect to the highest probability by updating its mean and covariance values. This method assigns each data point exactly to one operational region (i.e. engine mode) of the engine-propeller combinatory diagram. Therefore, the boundaries of the most frequent operating region of the engine-propeller combinator diagram can be determined.

The E-step is by considering a multivariate GMM initiated and denoted as [14]:

$$p_j(x;\mu_j,\Sigma_j)=\frac{1}{(2\pi)^{\frac{n}{2}}\left|\Sigma_j\right|^{\frac{1}{2}}}e^{-\frac{1}{2}(x-\mu_j)^T\Sigma_j^{-1}(x-\mu_j)} \tag{1}$$

where x is a n-dimensional input data vector and $p_j(x;\mu_j,\Sigma_j)$ is the probability density function (PDF) of the j-th multivariate Gaussian distribution (i.e. GMM) with, μ_j and Σ_j, the mean and covariance matrices, respectively. The probability of the i-th data point belongs to the j-th GMM is:

$$w_j^{(i)}=p\left(z^{(i)}=j\mid x^{(i)};\phi,\mu,\Sigma\right) \tag{2}$$

One should note that (2) calculates the "soft guess value" for the parameter, $z^{(i)}$. Considering the Bayes rule and (1), the posterior probability of the parameter, $z^{(i)}$ given the parameter, $x^{(i)}$, can be written as:

$$p\left(z^{(i)}=j\mid x^{(i)};\phi,\mu,\Sigma\right)=\frac{p\left(x^{(i)}\mid z^{(i)}=j;\mu,\Sigma\right)p\left(z^{(i)}=j;\phi\right)}{\sum_{l=1}^{k}p\left(x^{(i)}\mid z^{(i)}=l;\mu,\Sigma\right)p\left(z^{(i)}=l;\phi\right)} \tag{3}$$

where $p\left(z^{(i)}=j;\phi\right)$ is the prior probability of the j-th GMM and k is the number of GMMs. The equal prior probability is of each GMM initially assumed. One should note that (3) represents a multivariate Gaussian distribution with μ_j and Σ_j are the mean and covariance matrixes, respectively and the j-th GMM is a m-dimensional data vector. Hence, the equation for M-step is:

$$\phi_j=\frac{1}{m}\sum_{i=1}^{m}w_j^{(i)}$$

$$\mu_j = \sum_{i=1}^{m} w_j^{(i)} x^{(i)} \Big/ \sum_{i=1}^{m} w_j^{(i)}$$
$$\Sigma_j = \sum_{i=1}^{m} w_j^{(i)} \left(x^{(i)} - \mu_j\right)\left(x^{(i)} - \mu_j\right)^T \Big/ \sum_{i=1}^{m} w_j^{(i)} \tag{4}$$

This step is to update the respective GMM (i.e. data cluster) used by calculating the new mean and covariance values with respect to each data point. This iterative process should stop either at the end of the training data set or approximately equal prior to or posterior to mean and covariance values.

Mean and Covariance Values

The EM algorithm may converge to a local minima or saddle point in some situations during its iterative process. Therefore, the initial mean and covariance values selection should be appropriately; the initial mean and covariance values are from the previous statistical distributions approximated (see Figure 2). The estimated mean and variance values for each Gaussian distribution in Figure 2 are:

$$\mu_1 = \left[92.5(rpm) \quad \log(4500)(kW - \log scale)\right]$$

$$\mu_2 = \left[101(rpm) \quad \log(5500)(kW - \log scale)\right]$$

$$\mu_3 = \left[111(rpm) \quad \log(7300)(kW - \log scale)\right]$$

$$\Sigma_1 = \Sigma_2 = \Sigma_3 = \begin{bmatrix} 4.6106(rpm^2) & 0.0551(rpm.kW - \log scale) \\ 0.0551(rpm.kW - \log scale) & 0.0007(kW - \log scale)^2 \end{bmatrix} \tag{5}$$

These values are in the initial GMMs and presented in left plot of Figure 7 introduced (i.e. the combinator diagram). These initial GMMs are near the local maxima points to improve the performance of the EM algorithm estimated. The respective GMMs are as multivariate Gaussian distributions with the mean and covariance values in (5) represented. The respective contours of the multivariate Gaussian distributions (i.e. GMMs) are by ellipse in this figure denoted. Then, the EM algorithm is to update these GMMs executed with the respective data points and the results are in right plot of Figure 7 presented. The results show that the GMMs are converged to appropriate mean and covariance values under the EM algorithm. Therefore, these three regions are as the most frequent operating regions of the engine-propeller combinator diagram identified by considering the respective ship performance and navigation data. However, an overlay situation within two GMMs is also in these results observed. It is believed that such data overlay situations is possible to avoid by classifying the same data set in a higher dimensional space under additional ship performance and navigation parameters. The updated equation of mean and covariance values of the GMMs under the EM algorithm are:

$$\mu_1 = \left[92.8770(rpm) \quad 3.6515(kW - \log scale) \right]$$

$$\mu_2 = \left[101.5521(rpm) \quad 3.7549(kW - \log scale) \right]$$

$$\mu_3 = \left[110.4419(rpm) \quad 3.8609(kW - \log scale) \right]$$

$$\Sigma_1 = \begin{bmatrix} 8.8222(rpm^2) & 0.1159(rpm.kW - \log scale) \\ 0.1159(rpm.kW - \log scale) & 0.0021(kW - \log scale)^2 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 3.6268(rpm^2) & 0.0340(rpm.kW - \log scale) \\ 0.0340(rpm.kW - \log scale) & 0.0004(kW - \log scale)^2 \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} 1.4695(rpm^2) & 0.0030(rpm.kW - \log scale) \\ 0.0030(rpm.kW - \log scale) & 0.0001(kW - \log scale)^2 \end{bmatrix} \tag{6}$$

One should note that slight variations among the initial and final means values and considerable variations among the initial and final covariance values are in these results observed. These same observations are also under the results in Figure 7 noted.

CONCLUSIONS & FUTURE WORK

A selected data set of ship performance and navigation parameters is in this study used to capture the most frequent operating regions of main engine analyzed. This development approach categorizes as the data analytics under the engine-propeller combinator diagram for ship performance monitoring. The results show that the main engine is operated around three frequent regions in the combinator diagram that are identified by GMMs with the EM algorithm with the respective engine parameters: shaft speed, fuel consumption and ME power. These operating regions are considered as the basis for the SEEMP (i.e. the engine modes), where the respective vessel performance should be evaluated. The proposed data analytics can in the ship operation phase identify the optimal engine-propulsion operating conditions. Hence, engine and propeller operating regions are possible to select appropriately to reduce fuel consumption of vessels. Furthermore, development of advanced mathematical models of ship performance and navigation under these operating regions of marine engines will be the future work of this study. It is believed that such approach can overcome the current shipping industrial challenges under emission control based energy efficiency measures [15].

ACKNOWLEDGMENT

An initial version of this paper is presented at the 35th International Conference on Ocean, Offshore and Arctic Engineering, 2016, Busan, Korea. The authors would like to thank anonymous reviewers for their helpful and constructive comments that greatly contributed to improve the final version of the paper. High-resolution color images of this data analysis are available at site: http://bit.do/perera.

FUNDING

This work has been conducted under the project of "SFI Smart Maritime (237917/O30) - Norwegian Centre for improved energy-efficiency and reduced emissions from the maritime sector" that is partly funded by the Research Council of Norway.

REFERENCES

[1] IMO, 2009, *Resolution MEPC.1/Circ.683, Guidelines for the development of a ship energy efficiency management plan (SEEMP).*

[2] IMO, 2012, *Resolution MEPC.213(63), 2012 Guidelines for the development of a ship energy efficiency management plan (SEEMP).*

[3] IMO, 2009, *Resolution MEPC.1/Cric.684, Guidelines for the voluntary use of the ship energy efficiency operational indicator (EEOI).*

[4] Perera, L. P. and Mo, B., 2016, "Emission Control based Energy Efficiency Measures in Ship Operations," Journal of Applied Ocean Research, **60**, pp. 29-46.

[5] Perera, L. P. and B. Mo, 2016, "Data Compression of Ship Performance and Navigation Information under Deep Learning," *Proc. 35th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2016)*, Busan, Korea, (OMAE2016-54093).

[6] Perera, L. P., Mo, B., and Kristjansson, L. A., 2015, "Identification of Optimal Trim Configurations to improve Energy Efficiency in Ships," *Proc. 10th IFAC Conference on Manoeuvring and Control of Marine Craft (MCMC 2015)*, Copenhagen, Denmark, pp. 267-272.

[7] MAN Diesel & Turbo, 2011, "Basic principles of ship propulsion," Copenhagen, Denmark, pp. 24-32.

[8] Perera, L. P., Rodrigues, J. M., Pascoal, R., and Guedes Soares, C., 2012, "Development of an onboard decision support system for ship navigation under rough weather conditions," Sustainable Maritime Transportation and Exploitation of Sea Resources, E. Rizzuto & C. Guedes Soares (Eds.), Taylor & Francis Group, London, UK, pp. 837-844.

[9] Perera, L. P., Machado, M. M., Valland, A., and Manguinho, D. A. P., 2015, "System Reliability of Offshore Gas Turbine Engines with Erroneous Data Conditions," *Proc. 25th*

European Safety and Reliability Conference (ESREL 2015), Switzerland, September, pp. 1679-1688.

[10] Perera, L. P., and Mo, B., 2016, "Ship Speed Power Performance under Relative Wind Profiles," Maritime Engineering and Technology III , Guedes Soares & Santos (Eds.), vol. 1, Taylor & Francis Group, London, UK, pp. 133-141.

[11] Perera, L. P., and Mo, B., 2016, "Data Analytics for Capturing Marine Engine Operating Regions for Ship Performance Monitoring," *Proc. 35th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2016),* Busan, Korea, June, (OMAE2016-54168).

[12] Sun, S., Zhang, C., and Yu, G., 2006, "A Bayesian network approach to traffic flow forecasting," IEEE Transactions on Intelligent Transportation Systems, **7**(1), pp. 124-132.

[13] Moon, T. K., 1996, "The expectation-maximization algorithm," IEEE Signal Processing Magazine, **13** (6), pp. 47-60.

[14] Ng, A., 2016, "Mixtures of Gaussians and the EM algorithm," Lecture notes on Machine Learning, Stanford University, USA, (09/25/2016).

[15] Perera, L. P., and Mo, B., 2016, "Machine Intelligence for Energy Efficient Ships: A Big Data Solution," Maritime Engineering and Technology III , Guedes Soares & Santos (Eds.), vol. 1, Taylor & Francis Group, London, UK, pp. 143-150.

Figure Captions List

Table Caption List

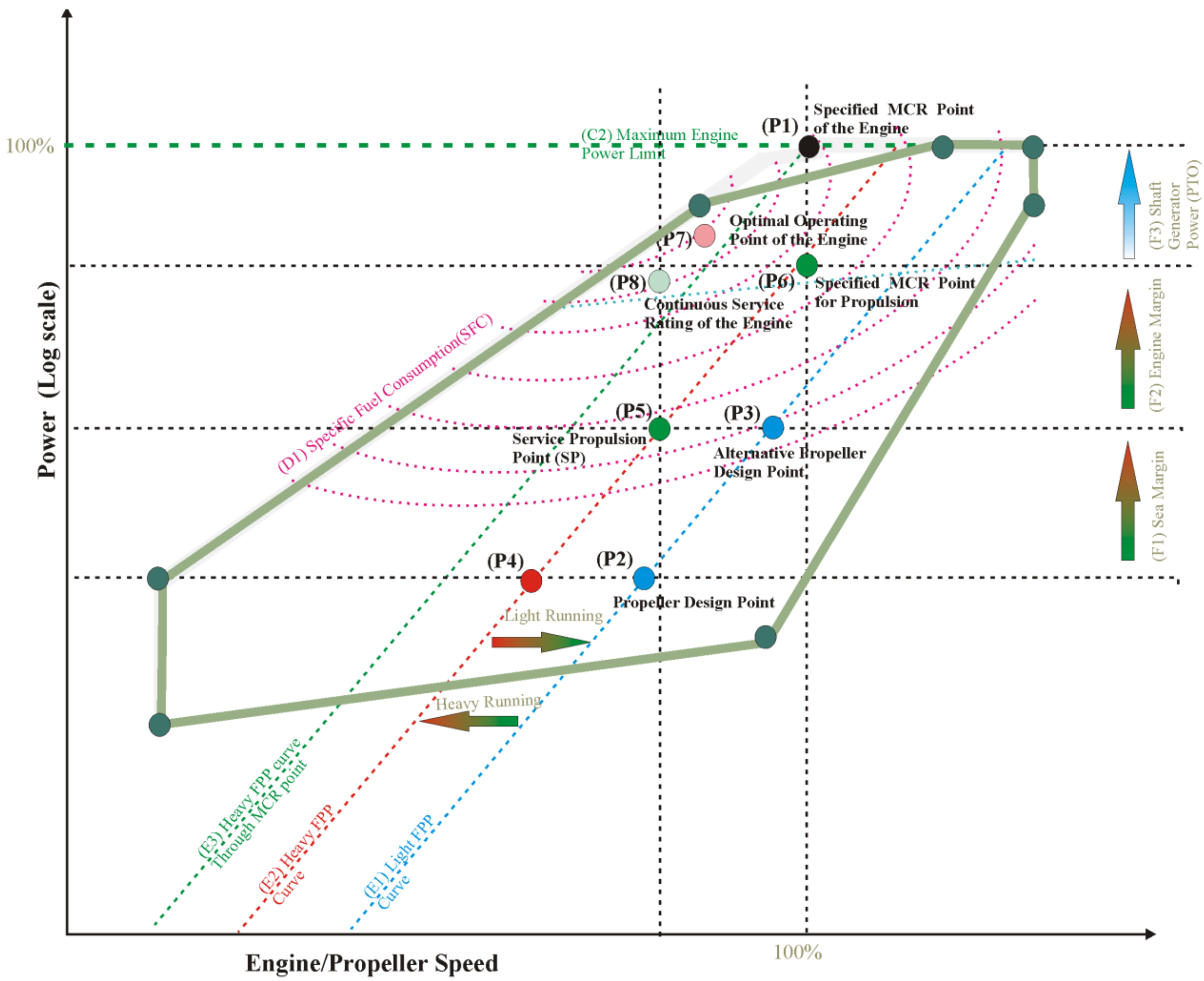

Figure 1. Simplified engine-propeller combinator diagram

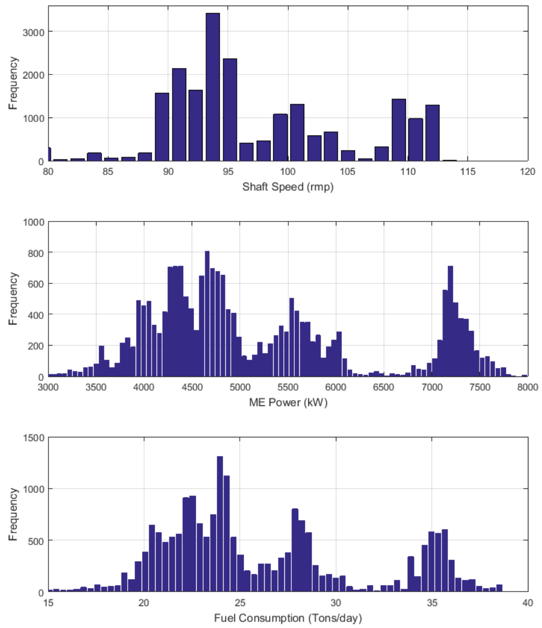

Figure 2. Statistical distributions of marine engine parameters

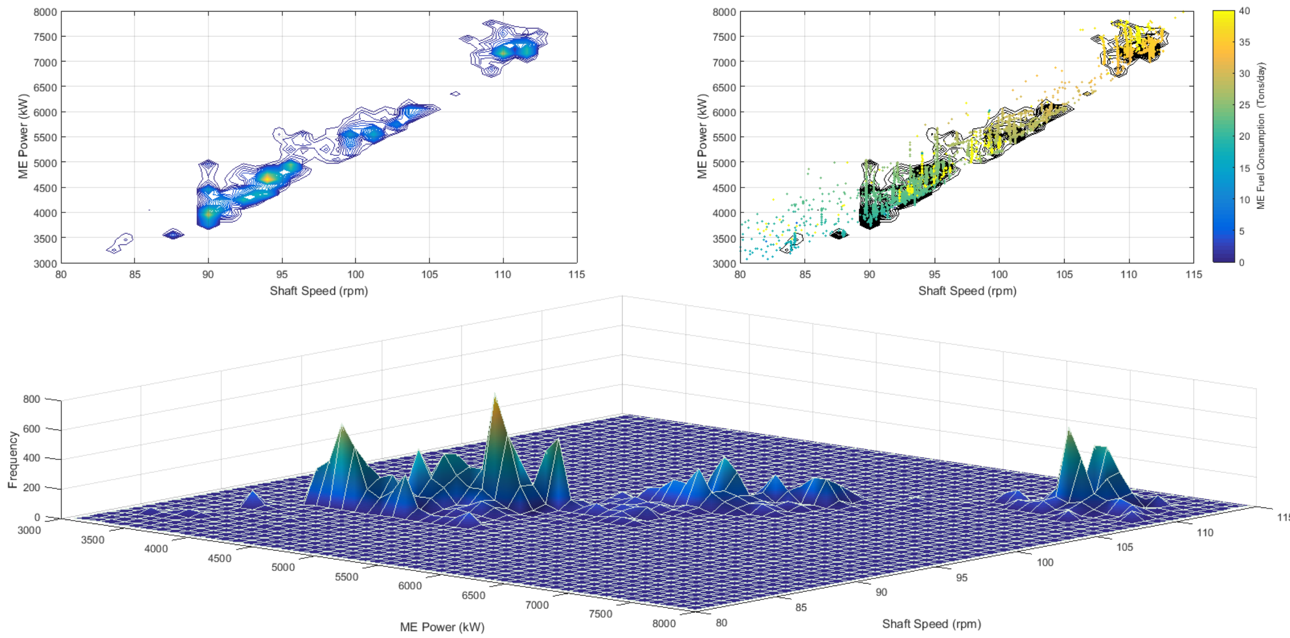

Figure 3. Engine operation regions: ME power vs. shaft speed

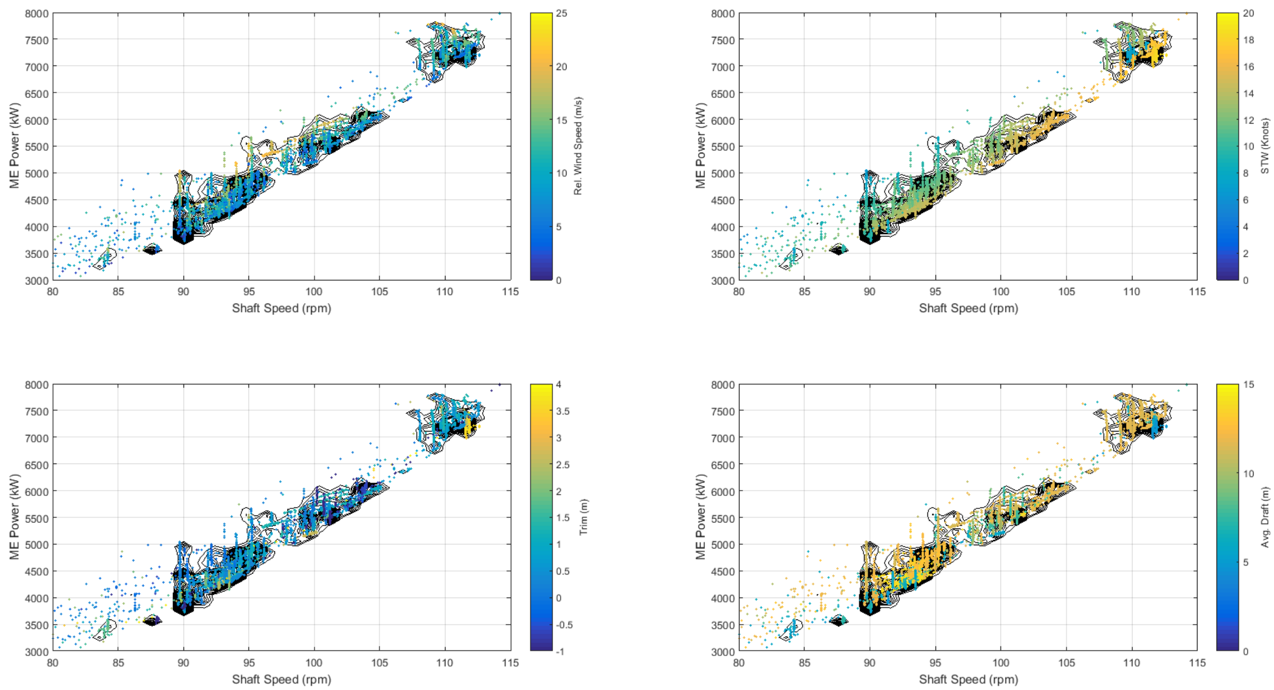

Figure 4. Engine operation region vs. rel. wind speed, STW, avg. draft and trim

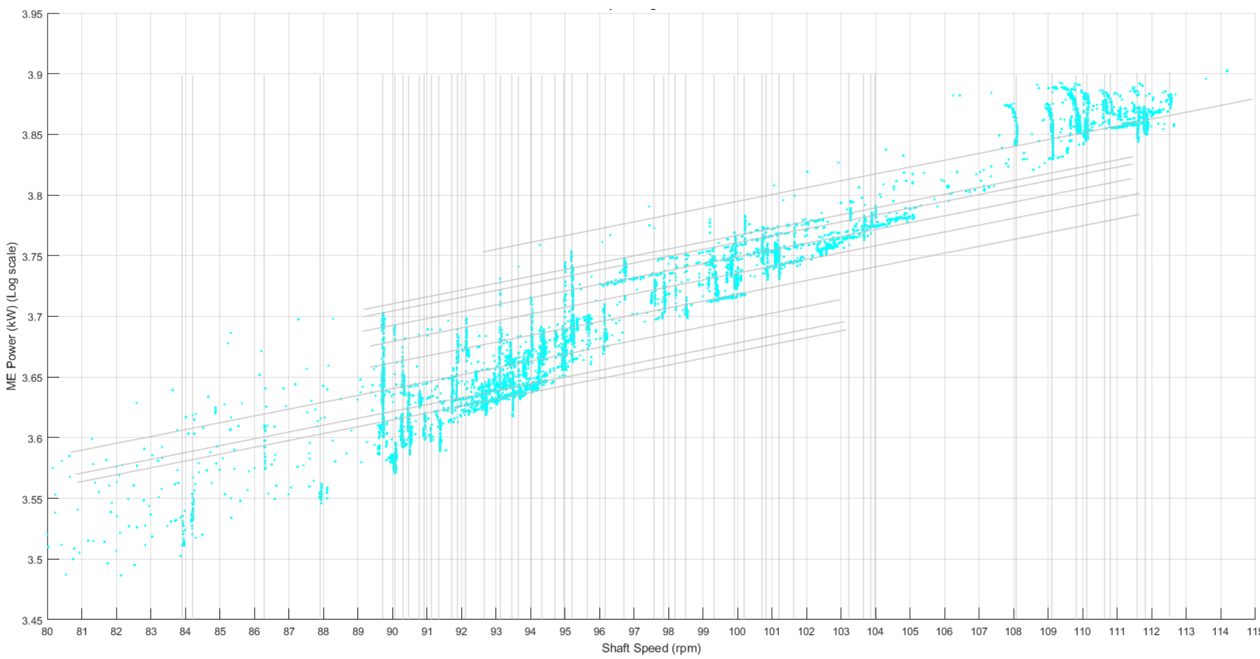

Figure 5. Engine propeller combinator diagram

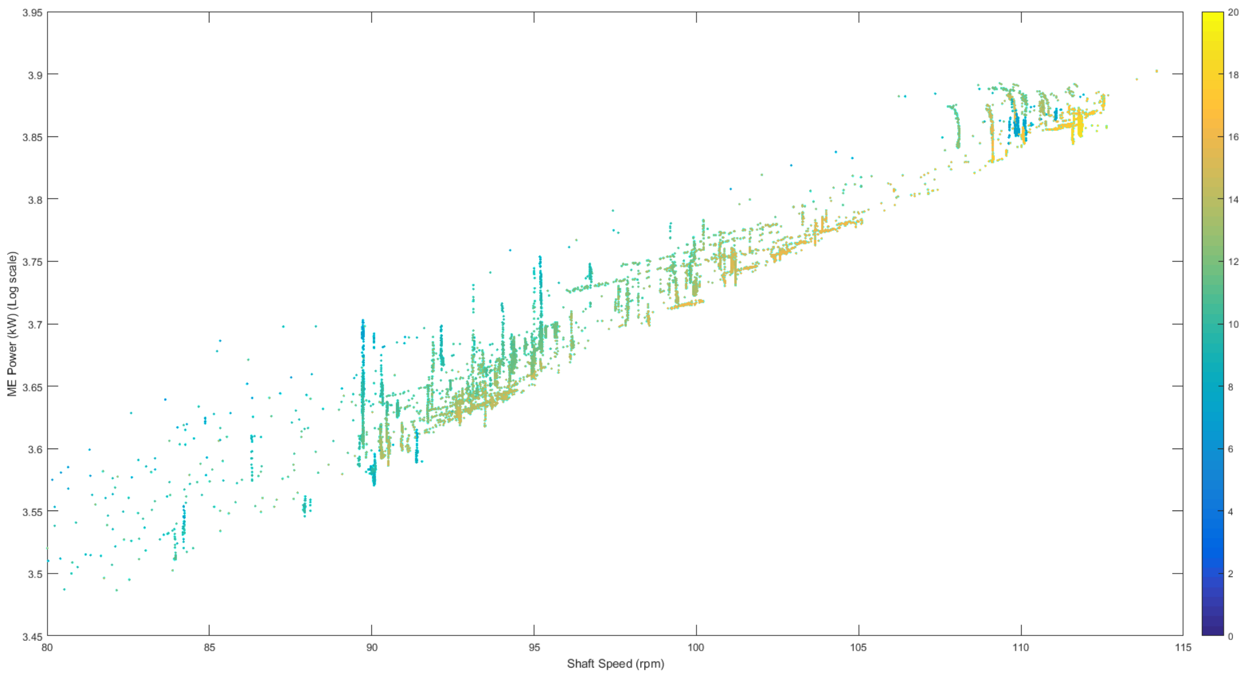

Figure 6. Engine propeller combinator diagram with STW

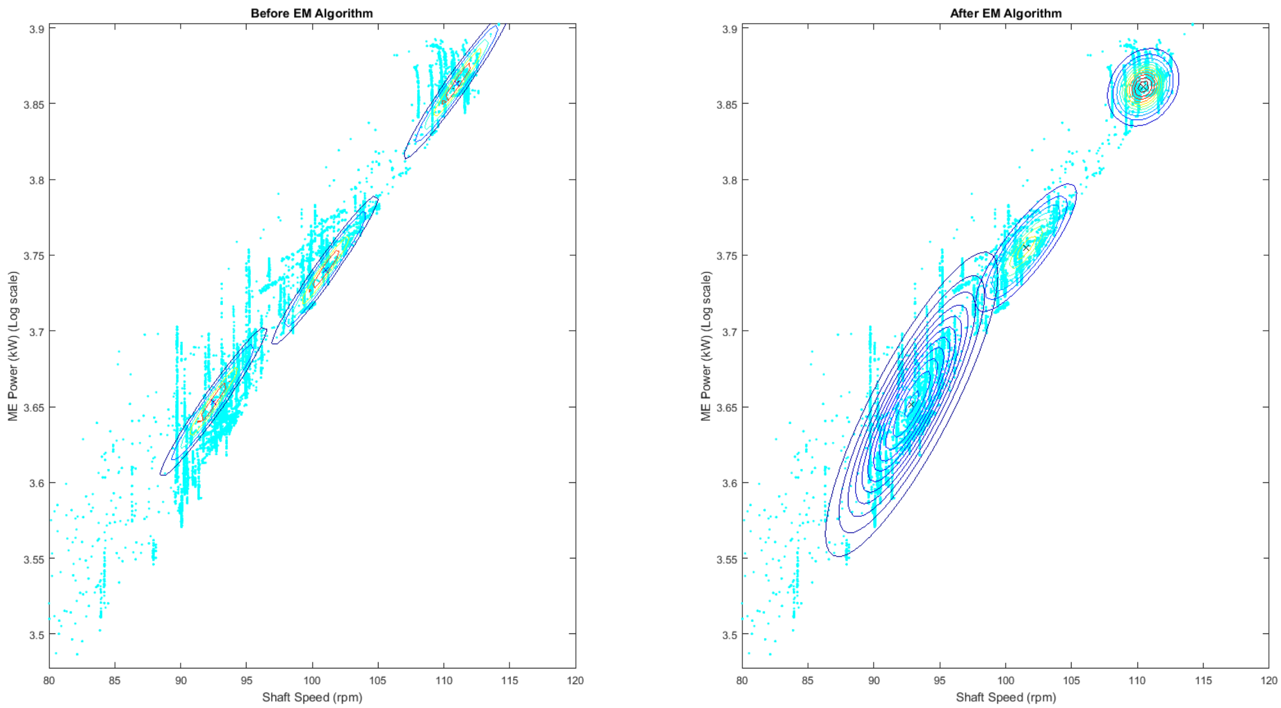

Figure 7. GMMs in the engine propeller combinator diagram

Parameter	**Description**
Ship Type	Bulk carrier
Ship length	225 (m)
Ship beam	32.29 (m)
Gross tonnage	38.889 (tons)
Deadweight (at max draft)	72.562 (tons).
Engine type	2 stroke main engine MCR of 7564 (kW) at 105 (rpm)
Propeller type	A fixed pitch propeller with diameter 6.20 (m) and 4 blades.

Table 1: Vessel particulars.