Ship fuel consumption monitoring and fault detection via partial least squares and control charts of navigation data

Capezza C. [1], *Coleman S.*[2], *Lepore A.*[1], *Palumbo B.*[1], *Vitiello L.*[1]
[1]University of Naples "Federico II", Naples, Italy
[2]Newcastle University, UK

Declarations of interest: none.

ABSTRACT

New regulations in the shipping sector aim to give greater transparency to operations and public access to CO_2 emissions data. EU regulation 2015/757 became mandatory in January 2018 and urges shipping companies to set up systems for daily monitoring, reporting and verification (MRV) of emissions for individual ships. Manual acquisition and handling of emissions data may be allowed (e.g. bunker fuel delivery note, bunker fuel tank monitoring), but is adversely affected by uncertainty due to human intervention and will eventually be unusable for monitoring purposes. However, the massive amounts of navigational data acquired by multi-sensor systems installed on-board modern ships have great potential to aid compliance with regulations but their use is hampered by the lack of effective analytical methods in maritime literature. This work demonstrates a statistical framework and automatic reporting system for fuel consumption monitoring that addresses the MRV requirements needed to comply with the regulations. The framework has been applied to the Grimaldi Group's Ro-Ro Pax cruise ships and is shown, in addition, to be capable of supporting fault detection as well as verifying CO_2 savings achieved after energy efficiency initiatives.

Keywords: MRV, CO_2 emission monitoring, vessel energy efficiency, partial least squares regression, Hotelling's T^2 control chart, squared prediction error control chart.

1. Introduction

1.1 Regulatory background

In recent years, the increase of greenhouse gas (GHG) emissions such as carbon dioxide (CO_2) has determined global warming and climate change. They are considered one of the biggest challenges of our time and prompt solutions have to be adopted to avoid severe consequences for society. According to the Kyoto Protocol, several institutions have focused their attention on this problem. The International Maritime Organization (IMO), through the Maritime Environmental Protection Committee (MEPC), has developed initiatives to reduce GHGs from ships, urging shipping companies to adopt a set of technical and operational measures to improve energy efficiency of ships not only during operation but also in the design phase (IMO, 2009a, 2009b). The Energy Efficiency Design Index (EEDI) (IMO, 2009b), and the Ship Energy Efficiency Management Plan (SEEMP) (IMO, 2009a), are the tools used to monitor these improvements. Alongside IMO guidelines, the Regulation EU 2015/757 of the European Parliament forces shipping companies, operating with their fleet in the EEA (European Economic Area) regardless of the flag state or port registry, to monitor and report all harmful emissions (Council of European Union, 2015) from 1 January 2018.

The basic measure adopted by shipping companies to reduce CO_2 emissions and fuel consumption is the sailing speed reduction, since it does not require any energy efficiency operation that has a cost. In fact, the ship's speed has also implications in terms of energy efficiency and costs for owners and ship operators. Small changes in speed can significantly improve the energy efficiency as well as the productivity and revenue of the ship (Smith et al., 2011). The problem of optimizing the ship speed on a route in order to minimize the total fuel cost while satisfying the calling time window constraints at the calling ports has been faced by Kim et al. (2016, 2014). Because of the well-known difficulties in directly measuring CO_2 emissions, the MRV regulation (Council of European Union, 2015) provides for indirect monitoring through the ship's fuel consumption. The calculation of CO_2 emissions can, in fact, be retrieved on the basis of the amount of fuel consumption, through the emission factor in accordance to Annex VI (European Parliament and of the Council, 2012).

1.2 Literature review

In the naval literature, the most common method used to estimate fuel consumption and then CO_2 emissions is the so-called speed-power curve. This curve is drawn by exploiting the univariate relationship between the engine power and the vessel speed (Lewis, 1988). However, despite its intuitive usage, this method is affected by large variability and may lead to poor predictions due to different sailing (e.g. trim, displacement, etc.) and weather conditions.

Several methods have been proposed to improve the estimate of speed-power curves, by exploiting any information from additional operational variables. In particular, Bialystocki and Konovessis (2016) firstly consider ship's draught and displacement, weather force and direction, hull and propeller roughness. Perera and Mo (2016) draw the empirical relationship between fuel consumption and the main operational variables using a graphical data analysis of performance and navigation parameters to support the management in analyzing the energy flow path. Petersen, Jacobsen, and Winther (2012) proposed a statistical method based on artificial neural networks and Gaussian Processes to predict fuel consumption through sailing and environmental conditions such as vessel's speed over ground, vessel's speed through water, trim, displacement, wind force and directions. Lu, Turan, and Boulougouris (2013) estimated the ship resistance considering as operational variables the ship type, draughts, speeds, encounter angles, sea states, fouling effect and engine degradation conditions. Meng, Du, and Wang (2016) proposed a method based on regression models to estimate the fuel consumption rate of container ships. Trodden et al (2015) analyzed fuel consumption and speed over the ground using a continuous data stream from a tug boat. Murphy et al (2012) used fuel consumption data and engine load from sea trials to investigate reduction in fuel consumption. Zaman et al. (2017) presented a statistical analysis to automatically detect the vessel operational modes (port, manoeuvring, sailing) based on sensor data acquired on board as a pre-cursor to modelling fuel consumption in transit mode.

Bocchetti et al. (Bocchetti et al., 2015, 2013) proposed a statistical method based on multiple linear regression to predict ship fuel consumption and build prediction intervals for each voyage considering the operational and sailing conditions including ship speed, sailed distance, wind speed, wind direction, cumulative docking time,

displacement, stabilizer fin operating time, and engine operation mode. In particular, this method exploits the massive amount of sensor data acquired on board of two Ro-Ro Pax ships that link three ports in the Mediterranean Sea.

1.3 The proposed approach

Unfortunately, most of the presented methods have strong limitations when applied to high-dimensional and correlated data, or they do not fully exploit all of the available information. New data acquisition technologies in fact have brought massive navigation data and a call for shipping management to adopt new methodologies to fully exploit them. For this purpose, an engineering approach based on partial least squares (PLS) regression is introduced to develop a model for ship fuel consumption prediction and monitoring based on the massive navigation data automatically acquired on-board of the modern ships. In particular, the proposed model is based on summary statistics of each voyage deduced from the sensor signals that relate to the actual navigation time and therefore, neglects manoeuvre time and stay in port at departure and at arrival. The fuel consumed in this phase is known to contribute more than the 90 percent to the total consumption.

The aim of the proposed approach is twofold: (*i*) statistical monitoring of ship fuel consumption (and thus CO_2 emissions) to support shipping management to identify anomalies and (*ii*) quantifying fuel consumption reduction (FCR) consequent to energy efficiency initiatives (EEIs) (hereinafter referred to also as *dry-dock* operations). The rest of the paper is organized as follows. A brief introduction on the acquisition data system and the variables used in this procedure is given in Section 2. Section 3 details the statistical approach and the monitoring tools utilized. Lastly, in Section 4 a real-case study is presented to illustrate the applicability and the effectiveness of the proposed approach.

2. Data

The navigation data are collected on board of four Ro-Ro Pax ships owned by Grimaldi Group that operate in the Mediterranean Sea. Ships, port names and dates are intentionally omitted for confidentiality reasons as well as actual numeric values but axis scales in the figures are left unchanged. However, even if we have currently been implementing the model on all the four Ro-Ro Pax ships mentioned above, for reasons

of brevity, in what follows we show only the most relevant results achieved for two ships (hereinafter referred to as *Ship 1* and *Ship 2*).

Each ship is equipped with a sensor network and a data acquisition device (DAQ). Every five minutes the DAQ device collects the values of a set of physical variables used as predictors, the complete set of physical variables of engineering interest used as predictor variables to monitor the ship operating conditions, ship fuel consumption, and CO_2 emissions is illustrated in Table 1. Further information on predictor variables and their descriptions are given in the following sub-section.

Table 1 should be arranged here or below.

2.1 Variable Definition

As previously stated, each variable summary statistic is available only at the end of the actual navigation time (h), that is defined as the time interval between the *Finished with Engine* order at departure port and *Stand by Engine* order at arrival port (IMO, 2000). The actual navigation time is calculated on the data and time UTC Coordinated Universal Time acquired through the Saab R4 GPS Navigation Sensor.

The Speed Over Ground (V) and the variance of SOG (σ_V^2) represent respectively the mean value and the variance value of the averages obtained from the 5 minutes interval observations. In particular, the SOG is obtained as

$$V = \frac{M}{h}, \tag{1}$$

where h is the actual navigation time and M is the distance travelled, calculated from the latitude and longitude data collected by the Saab R4 GPS Navigation Sensor. The distance travelled is calculated through the haversine formula according to (Veness, 2007) as

$$M = r\theta, \tag{2}$$

where r is the radius of the Earth, i.e. 6378.14 km and θ is defined as an angular distance in radians. The angle θ is calculated via the haversine formula as follows

$$hav\theta = hav\Delta\varphi + cos\varphi_A cos\varphi_B hav\Delta L, \tag{3}$$

where $\Delta\varphi$ is the difference in latitudes, ΔL is the difference in longitudes, while φ_A and φ_B are the latitude values at point A and point B, respectively, and $hav\theta = \sin^2(\theta/2)$ is the haversine function (Inman, 1849).

Specific resistance tests carried out in a towing tank within a speed range of 20 to 25 kn show that, for the considered ship type, the hydrodynamic resistance is proportional to the third power of SOG. Therefore, the proposed approach adopts a SOG that is cubed in order to obtain V^3.

The wind speed and direction are considered through the head (W_H), side (W_S) and following (W_F) wind components. According to (ITTC, 2017), Figure 1 shows the sign convention. The head wind component is calculated as

$$\tilde{W}_H = \begin{cases} 0 & \text{if } 90° \le \psi_{WT} \le 270° \\ V_{WT}\cos\psi_{WT} & \text{otherwise} \end{cases}, \tag{4}$$

where the V_{WT} is the true wind speed and Ψ_{WT} the difference between the true wind angle in earth system (B_{WT}) and course over ground (COG) averaged by the DAQ device every 5 minutes. V_{WT} and B_{WT} are automatically calculated by a Thies Clime anemometer (Adolf Thies GmbH & Co. KG, Gottingen, Deutschland) based on the sensor measurements of the relative wind speed V_{WT} and direction Ψ_{WT}, as well as COG and SOG data. The wind speed sensor has an accuracy of ±2.5% and a resolution of 0.05 m, while the wind direction sensor has an accuracy of ±2.5% and a resolution of 2.5°. The side wind component W_S is defined once a voyage is completed as the mean of

$$\tilde{W}_S = \left| V_{WT}\sin\psi_{WT} \right|, \tag{5}$$

while the following wind component W_F is defined as the mean value of

$$\tilde{W}_F = \begin{cases} -V_{WT}\cos\psi_{WT} & \text{if } 90° \le \psi_{WT} \le 270° \\ 0 & \text{otherwise} \end{cases}. \tag{6}$$

Figure 1 should be arranged here or below.

The engine variables are port and starboard shaft generator power variables SG_P and SG_S and they allow taking into account the different modes of navigation (constant and combinator mode).

The Power difference between two propeller shafts ΔP allows discovering anomalies or malfunctions in the main engines, while the Power difference between two shaft generator powers ΔSG allows discovering if one of the two shaft generators is out of order.

The Displacement variable Δ is defined as the mean value of the two displacements obtained, respectively, at departure and arrival port. In particular, each of these two displacement values are derived from the hydrostatic data based on draughts at amidships and trim. The former is obtained by averaging the portboard and the starboard draught in the amidships section at the departure and arrival port, i.e. T_{PD}, T_{SD}, T_{PA}, and T_{SA}; the latter in accordance with (ITTC, 2008) is defined as the mean value of difference between the draught fore and the draught aft at the arrival port, i.e. T_{FA} and T_{AA} . The draught variables T_{FD}, T_{AD}, T_{PD}, T_{SD}, T_{FA}, T_{AA}, T_{PA}, and T_{SA} do not refer to the entire voyage, but each variable is measured both at departure and at arrival ports by four submersible transmitters located at fore and aft perpendiculars, and at port and starboard amidships sections. These data are acquired by four Vegawell 52 draught gauges (pressure transmitters; VEGA, Schiltac, Deutschland), each with a maximum deviation of 0.2%. These measurements are collected in port when SOG is less than 0.3 Knots because this sensor acquires the hydrostatic pressure. At high speed these measurements are affected by errors.

Departure and arrival trim ($Trim_D$ and $Trim_A$) are obtained through the inclinometer measurements and the geometric features of the ship.

2.2 Hourly Fuel consumption and CO_2 emission calculation

The response variable object of this study is the average fuel consumption per hour for each voyage (i.e. the ratio between fuel consumed and actual navigation time in hours for each voyage) during navigation

$$Y = \frac{Q}{h} \ , \tag{7}$$

where Q is the total fuel consumption for the voyage during the actual navigation time and h is the navigation time in hours. A brief overview of the ship engine room layout is presented to identify how Q is calculated. All the cruise ships monitored in this paper have two engine sets, each with two Wartsila 12V46D main engines for propulsion

with a variable pitch propeller and a Marelli shaft generator (keyed on a gearbox) for electric power; Figure 2 outlines the Engine room layout.

Figure 2 should be arranged here or below.

As detailed by Bocchetti et al. (2015), on the *j*-th engine set (*j*=1, 2) the DAQ device collects the thrust power P_j^T on the shaft propeller and the electrical power P_j^E on the shaft generator. The powers P_j^T and P_j^E are the only measurements available for calculating the main engine power. Note that when the engine operation mode is combinator, the shaft generator is necessarily powered off ($P_j^E = 0$). In this case, the electrical power is supplied by three diesel generators, which are intentionally not considered in the following fuel consumption calculation.

The actual fuel consumption *Q* related to the main engines is calculated through the following relation

$$Q = \sum_{i=1}^{2} \sum_{j=1}^{2} P_{ij} h_{ij} SFOC_{ij} \ , \tag{8}$$

where h_{ij} is the number of running hours of the *i*-th engine of the *j*-th engine set (with *i*,*j*=1,2), $SFOC_{ij}$ is the specific fuel oil consumption of the *i*-th engine of the *j*-th engine set. Then, the power P_{ij} of the *i*-th main engine of the *j*-th engine set can be calculated as follows

$$P_{ij} = \begin{cases} 0 & \text{if } x_{ij} = 0 \\ \dfrac{x_{ij}}{\sum_{i=1}^{2} x_{ij}} P_j & \text{otherwise} \end{cases}, \tag{9}$$

where x_{ij} assumes the value 0 if the *i*-th main engine of the *j*-th engine set is powered off and 1 otherwise. For each voyage, the main engine power P_j of the *j*-th engine set is calculated as the mean value of

$$P_j = \frac{P_j^E}{\eta_j^e \eta_j^m} + \frac{P_j^T}{\eta_j^m} \ , \tag{10}$$

where η_j^m and η_j^e are the gearbox mechanical efficiency and the shaft generator electrical efficiency, respectively.

The proposed method can utilize CO_2 emissions as response variable, in particular, according to Annex I of MRV regulation (Council of European Union, 2015), the calculation of CO_2 emissions can be performed exploiting the amount of ship fuel consumption through the following formula

$$CO_2 \text{emission} = \text{fuel consumption} \times \text{emission factor}$$

For each fuel type, a different value of the emission factor is available, according to the Intergovernmental Panel for Climate Change (IPCC) as reported in Annex VI (European Parliament and of the Council, 2012).

2.3 Timeline and Maintenance Intervals

As already stated, for reasons of brevity, we show the most relevant results achieved for two of the four Ro-Ro Pax ships owned by Grimaldi Group. In particular, the application of the proposed approach for the on-line monitoring of fuel consumption and fault detection (aim (*i*)) is illustrated by means of data acquired on *Ship 1*; whereas, data acquired on *Ship 2* are used to show the capability of the proposed approach to assess and quantify the fuel consumption reduction related to a dry-dock operation (aim (*ii*)). In what follows, the data used to estimate the model are referred to as *calibration dataset*; whereas those used for aims (*i*) and (*ii*) are referred to as *monitoring dataset*. Accordingly, for *Ship 2* the calibration and monitoring datasets refer to data collected before and after EEI, respectively. The periods to which calibration and monitoring datasets refer, as well as those indicating dry-dock operations, are outlined in Figure 3 and Figure 4 for *Ship 1* and *Ship 2*, respectively.

For the sake of completeness, for *Ship 1*, the calibration and the monitoring period include 606 and 720 voyages, respectively; for *Ship 2* the calibration (before EEI) and the monitoring (after EEI) datasets contain 329 and 462 voyages, respectively. In particular, for *Ship 1*, 11 months' worth of data collected right after EEI operation have been certified as reference data for the model calibration, in the extent of capturing all the typical operating conditions; whereas, for *Ship 2*, 10 months' worth of data between two dry-dock operations have been certified as reference data for the model calibration,

Anomalous voyages have been identified by proper statistical procedures and are left out in the reference dataset only if being confirmed as exceptional by technical engineers.

Figure 3 and Figure 4 should be arranged here or below

3. The Statistical Approach and Monitoring Tools

In this paper, PLS regression is used to evaluate the parameters of the statistical model for predicting and monitoring ship fuel consumption and emissions of CO_2. The choice of the PLS in place of e.g., multiple linear regression (Erto et al., 2015) has great potential of supporting the management to handle the great amount of data collected on board of modern ships that are usually noisy and strongly correlated. The residual left by the PLS model are also naturally prone to be monitored at each new voyage through *prediction error control chart*, whereas the predictor variables are monitored through the *Hotelling's* T^2 and SPEx control chart, as detailed below. When a point falls outside the upper control limit of at least one of the latter control charts, a possible problem may have occurred. The management is then urged to further investigate physical variables that have caused the out-of-control condition by exploring the corresponding *contribution plot* (MacGregor and Kourti, 1995).

From a mathematical point of view, the two statistics monitor different anomalies that may occur during voyages. In particular, a value exceeding the control limits of the Hotelling's T^2 chart indicates that the corresponding observation presents extreme values in one or more physical variables, but plausibly maintains the same correlation structure as in the reference dataset (high variability *inside* the PLS model). In opposition, values exceeding the control limits in the SPE_x control chart are related to observations that have a different structure with respect to the reference data (outside the PLS model).

In Figure 5 the main steps of the statistical procedure proposed for monitoring of fuel consumption and diagnosis of faults is outlined. The first step is to set up the monitoring control chart of the T^2 statistic and identify the anomalous voyages for which the statistic falls outside the upper control limit. For each of these voyages, in order to identify the physical variables that have determined the largest value, the

contribution plot to T^2 is built. The second step regards the SPE_x control chart. Similarly, to the T^2 statistic, for all voyages with the SPE_x statistic falling above the upper control limit, a contribution plot to the SPE_x statistic is performed. Finally, the prediction error control chart is built up. More information is given in the next section, which shows the procedure applied operatively to a real case study from a Ro-Ro Pax ship.

Figure 5 should be arranged here or below.

3.1 The Hotelling T^2 chart

Note that the detailed explanation of the statistical considerations about the monitoring statistics introduced in the previous subsection are not in the scope of this article, readers are referred to (Nomikos and MacGregor, 1995a) for more discussion on this topic. When dealing with a new voyage to be monitored, let us denote by $\mathbf{x}_{new}$ the vector of observations of the predictor variables from which the (multivariate) sample mean of the reference data is subtracted and with $\mathbf{t}_{new}$ the vector of the corresponding observations of the latent variables (or scores) for this voyage. Moreover, let us denote with $\mathbf{S}$ the sample covariance matrix of the scores of the reference observations.

The Hotelling T^2 chart is a monitoring chart reporting for a single voyage to be monitored the relative statistic T^2, which is calculated as

$$T^2 = \frac{\mathbf{t}_{new}^T \mathbf{S}^{-1} \mathbf{t}_{new} N(N-R)}{R(N^2-1)}, \tag{11}$$

Where N is the number of reference observations and R is the number of latent variables included in the model. Figure 6 shows the graphical meaning of T^2 statistics. Since T^2 has the F-distribution with R and N-R degrees of freedom, the upper control limit T^2_{limit} is defined as follows

$$T^2_{\text{limit}} = F_{R,N-R,\alpha}, \tag{12}$$

Where α=0.01 is the significance level. When the T^2 statistics for a new voyage falls above T^2_{limit}, a possible problem in a physical variable may have occurred. In particular, the value of one or more physical variables is unusual and gives a high contribution to the T^2 statistics that have determined an out of control signal in the T^2 statistic. It is possible to identify the physical variables that have the highest

contribution to the out-of-control signal by calculating the contribution to the T^2 statistic according to (MacGregor and Kourti, 1995), as the elements of the vector

$$\text{Contribution}_{T^2} = \mathbf{t}_{\text{new}}\mathbf{p}_j^T, \tag{13}$$

Where $\mathbf{p}_j$ is the *j*-th column of the loading matrix $\mathbf{P}$. Graphically, the plot of the contributions to T^2 is a bar plot, each bar displaying the corresponding physical variable used in the statistical model. The physical variables that have determined the largest contribution to the T^2 statistic are identified by the highest bars in the contribution plot. On the y axis, the values of the physical variables contribution are reported.

3.2 The SPE_x chart

For each voyage, the statistic reported in the SPE_x control chart is as follows (MacGregor and Kourti, 1995)

$$SPE_x = (\mathbf{x}_{new} - \hat{\mathbf{x}}_{new})^T(\mathbf{x}_{new} - \hat{\mathbf{x}}_{new}), \tag{14}$$

where $\hat{\mathbf{x}}_{new} = \mathbf{t}_{\text{new}}\mathbf{P}^{\text{T}}$ is the prediction of $\mathbf{x}_{new}$ based on the PLS latent variable model. Accordingly, the upper control limit is defined as

$$SPE_{\text{sup}} = \theta_1\left[\frac{c_\alpha\sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0(h_0-1)}{\theta_1^2} + 1\right]^{\frac{1}{h_0}}. \tag{15}$$

A point that falls outside of this limit may indicate that a variable has assumed an atypical value with respect to the other variables, and an exceptional cause may have occurred with respect to the baseline used to calibrate the model, as shown in Figure 7. As for the T^2 statistic, the contribution to the SPE_x

$$\text{Contribution}_{SPE} = \mathbf{x}_{new} - \hat{\mathbf{x}}_{new} \tag{16}$$

can be plotted to have indications on physical variables that may have concurred to the out-of-control.

Variable contributions to T^2 and SPE_x statistics are such that large contributions in absolute value lead to large values of the control statistics. Then, when an out of control signal is detected in one of the two control charts, one can identify the responsible variables by looking at the larger absolute contribution values.

3.3 The prediction error control chart

As the residuals left by the PLS model are also naturally prone to be monitored at each new voyage a *prediction error control chart* can be defined by plotting the predicted hourly fuel consumption $\hat{y}_{new}$ and the $100\cdot(1-\alpha)$ prediction interval calculated as in (Nomikos and MacGregor, 1995b)

$$\hat{y}_{new} \mp t_{\alpha/2;n-R-1}\sqrt{\frac{\mathbf{e}^{\mathrm{T}}\mathbf{e}}{n-R-1}\left[1+\mathbf{t}_{new}^{\mathrm{T}}\left(\mathbf{T}^{\mathrm{T}}\mathbf{T}\right)^{-1}\mathbf{t}_{new}\right]}. \quad (17)$$

The matrix $\mathbf{T}$ is the matrix of the scores based on the reference data, $\mathbf{e}$ is the vector of the prediction error of the response variable for the reference observations, and $t_{\alpha/2;n-R-1}$ is the $100\cdot\alpha/2$ percentile of a Student's distribution with *n-R-1* degrees of freedom. Note that from an engineering point of view, the use of the prediction error control chart is discouraged when the T^2 and SPE_x control charts signal an out of control, since this can cause the problem of *extrapolation*, i.e., using the estimator $\hat{y}_{new}$ in Equation (17) beyond the boundary of the predictor space (Montgomery et al., 2012). On the other hand, anomalies detected only by the prediction error control chart cannot be addressed to any of the monitored variables and therefore, they plausibly pertain to a change among factors not included in the set of predictor variables. Moreover, the usefulness of the prediction error control chart is the possibility to detect anomalous trends of the response variable over subsequent voyages or shifts from zero of the prediction error mean, which are not observable by neither the T^2 nor SPE_x control charts.

4. Real-case study

However, even if we have currently been implementing the model on all the four Ro-Ro Pax ships mentioned above, for reasons of brevity, in what follows we show only the most relevant results achieved for two ships (hereinafter referred to as *Ship 1* and *Ship 2*). In particular, Section 4.1 show the capability of this procedure to support management to identify anomalous fuel consumption (and thus CO_2 emissions) and the physical variable (prognosis of fault) that give the highest contribution to the out-of-control are identified for *Ship 1*. whereas, Section 4.2 illustrates how the approach

is able to quantifying FCR consequent to a silicone foul release coating of the hull of *Ship 2*.

4.1 Prognosis of faults

In this section, for illustrative purposes a relevant case study is presented to underline the ability of the proposed procedure to discover anomalous voyages and to support management in making suitable decisions to solve the problem that occurred on *Ship 1*. Note that each voyage is identified with a unique voyage number (VN).

In this case, a monitoring window from VN 2034 to 2053 of *Ship 1* is considered. With respect to this monitoring window, Figure 6, Figure 7 and Figure 8 show the Hotelling T^2 control chart, the SPE_x control chart, and the prediction error control chart, respectively. In particular, VN 2035 exceeds the upper control limit in all the three charts:

Figure 6, Figure 7 and Figure 8 should be arranged here or below.

As previously explained, according to the procedure displayed in the flow chart of Figure 5, when an observation falls outside the upper control limits in both the T^2 and SPE_x charts, to further investigate the cause that occurred, the contribution plots to T^2 and SPE_x are produced (Figure 9 and Figure 10).

Figure 9 and Figure 10 should be arranged here or below.

Draughts and displacement for voyage 2035 display the largest contribution to the T^2 statistic (Figure 9), which support technical investigations in detecting an error signal in the draft gauge sensor. The large contribution to SPE_x of the departure trim in Figure 10 is due to an unusual value of draught variables. In fact, trim and draught variables are related and an error signal in the draft gauge sensor turn into an error in the trim value accordingly.

Referring to Figure 6, for VN 2036 the Hotelling T^2 statistic falls outside the control limit, therefore the respective contribution plot is reported in Figure 11.

Figure 11 should be arranged here or below.

The physical variable responsible of the out of control in the Hotelling T^2 control chart is ΔP, i.e. the power difference between shaft propellers; further investigations revealed that the ship sailed with three engines in service during one hour of the total sailing time.

VN 2041 is out of control both in the SPE_x control chart (Figure 7), and in the prediction error control chart (Figure 8). For this voyage, a contribution plot to SPE_x is built in Figure 12.

Figure 12 should be arranged here or below.

The physical variable that gives the highest contribution of SPE_x is the draught in fore at departure. The cause is an error in the draft gauge sensor.

Instead, from VN 2047 to 2049 both the monitored statistics T^2 and SPE_x are in control (Figure 6 and Figure 7), while the prediction error control chart (Figure 8) underlines for these voyages an extra-consumption of fuel. In this case, contribution plots are not needed because the statistics T^2 and SPE_x do not exceed their upper control limits. The causes of extra-consumption are to be detected outside the variables considered in the statistical model.

For graphical reasons, Figures 6 through 8 show only 20 relevant voyages (from 2034 to 2053) out of the 720 included in the monitoring dataset of *Ship 2*.

4.2 Energy efficiency quantification

The statistical approach presented is also able to quantify the energy saving after an energy efficiency improvement (EEI) operation. In particular, the fuel consumption reduction and therefore the CO_2 emission reduction can be quantified through the saving z_i of a new voyage i defined as

$$z_i = (\hat{y}_i - y_i)h_i\,, \tag{18}$$

where $\hat{y}_i$ is hourly fuel consumption predicted for the new voyage, y_i is the actual hourly fuel consumption for the same voyage and h_i is the sailing time. As consequence, the quantity $\hat{y}_i - y_i$ represents the hourly saving after the EEI operation.

To quantify the total fuel consumption reduction (FCR) after an EEI operation, the set voyages to be considered refers to the monitoring period that immediately follows the

EEI operation. Only voyages that are in control in both Hotelling T^2 and SPE_x control chart are considered.

The FCR can be calculated as

$$\mathrm{FCR} = \sum_{i}^{N_{EEI}} z_i \,, \tag{19}$$

where N_{EEI} is the number of new voyages considered after the EEI to evaluate its effectiveness.

A relevant case study follows. Considering the *Ship 2*, the EEI operation that was set up for this ship is the installation of silicone foul release coating on the wetted surface of the hull. This is a non-toxic hull paint based on fluoropolymer and siloxane (silicone polymer) coating. The main property of this class of painting is its "non-stick" surface that prevents the attachment of marine organisms to the hull, therefore it avoids an increase in the resistance, which requires additional power and consequently additional fuel consumption (and CO_2 emissions) to maintain the same vessel speed. Setting up this kind of EEI operation, shipping management is able to determine the actual improvement in shipping performance as well as economic and environmental savings consequent to reduction in fuel consumption and CO_2 emissions. One year's worth of data of voyages after the dry-dock operation (VN 427 to 888), from 15 March 2016 to 14 March 2017 are monitored to quantify the FCR.

Figure 13 should be arranged here or below

For graphical reasons, Figure 13 shows the prediction error control chart for voyages 427—476 of Ship 2 to illustrate the improvement and the consequent saving gained after the installation of silicone foul release coating only for the first month following this operation. Note that the actual hourly fuel consumption is lower than the prediction hourly fuel consumption for all the voyages, underlying that the saving has been effectively obtained thanks to the EEI. Moreover, voyages displayed without their relative prediction error control chart are those outside upper control limit in Hotelling T^2 and/or SPE_x control chart, therefore according to (Nomikos and MacGregor, 1995b) their prediction interval is not calculated and not considered for the quantification of

FCR. In particular, the percentage of FCR quantified on one year of observations is 9.25%. This is a profitable result for the shipping management.

5. Conclusion

The shipping industry is facing a new regulatory regime that aims to give public access to CO_2 emissions data and this challenge can be addressed by making better use of the massive amounts of sensor data now available. The International Maritime Organization—through the Energy Efficiency Design Index and the Ship Energy Efficiency Management Plan—and the European Union—through the application of the EU regulation 2015/757, which is mandatory from January 2018—urge shipping companies to set up a system for daily monitoring of emissions from each ship. Engineers in the naval sector traditionally rely on deterministic relationships among the physical variables of interest in order to make decisions. Moreover, even if manual acquisition of emission data is allowed by all the international regulations (e.g. bunker fuel delivery note, bunker fuel tank monitoring), it is yet affected by uncertainty due to human intervention and thus will be eventually unusable for monitoring purposes. On the other hand, the massive amount of navigation data acquired by multi-sensor systems installed on-board of modern ships have a great potential to naturally comply with those regulations but are hampered by the lack of effective methods in the maritime literature. Despite this advantage, many shipping companies limit their analyses to only calculating simple summary statistics at each voyage. The range of tools needs to be extended to include statistical methods for detecting patterns in data. The main contribution of this work is to provide statistical tools that effectively support the management. Without those statistical tools, the management cannot easily investigate and diagnose faults responsible for increase in fuel consumption (and CO_2 emissions). The proposed statistical approach is therefore able to support managerial decision making by setting up suitable actions to improve ship performance as well as to quantify consumption/emission savings after energy efficiency improvement operations (e.g., hull form optimization, hull cleaning and propeller polishing, ultra-smooth coating, engine maintenance operation, propulsion and power plant efficiency improvement). As is known, this is particularly profitable for shipping companies in order to claim for carbon credit.

The proposed statistical approach is not only useful to maritime engineers and shipping companies, but also to the international organizations responsible to adopt regulations related to the CO_2 emission monitoring problem. In fact, the same techniques can be adopted to verify that shipping companies satisfy the regulatory requirements. Currently, however, these statistical tools are difficult to implement and spread among shipping companies because such organizations are not ready to be forced to use automatic systems for the statistical analysis of emission data in order to verify that CO2 emissions are coherent with the ship characteristic. The purpose of this paper is to describe the methodology and demonstrate its applicability. Ongoing implementation is expected to yield further evidence of the methodology's benefits and usability.

Acknowledgements

The authors are grateful to the Energy Saving Technical Department engineers Dario Bocchetti, Andrea D'Ambra and Rosa Di Matteo of the Grimaldi Group shipping company. This work is under the agreement of collaboration on big data in shipping for monitoring and benchmarking fuel efficiency and CO_2 emissions, between the Industrial Statistics Research Unit (ISRU) of the Newcastle University (United Kingdom), the Department of Industrial Engineering (DII) of the University of Naples Federico II (Italy), and the Process System Engineering Group (GEPSI) of the Chemical Process Engineering and Forest Products Research Center (CIEPQPF), Department of Chemical Engineering, University of Coimbra (Portugal).

References

Bialystocki, N., Konovessis, D., 2016. On the estimation of ship's fuel consumption and speed curve: A statistical approach. J. Ocean Eng. Sci. 1, 157–166. https://doi.org/10.1016/j.joes.2016.02.001

Bocchetti, D., Lepore, A., Palumbo, B., Vitiello, L., 2015. A statistical approach to ship fuel consumption monitoring. J. Sh. Res. 59. https://doi.org/10.5957/JOSR.59.3.150012

Bocchetti, D., Lepore, A., Palumbo, B., Vitiello, L., 2013. A statistical control of the ship fuel consumption, in: RINA, Royal Institution of Naval Architects - Design

and Operation of Passenger Ships 2013.

Council of European Union, 2015. Regulation (EU) 2015/757 of the European Parliament and of the Council of 29 April 2015 on the monitoring, reporting and verification of carbon dioxide emissions from maritime transport, and amending directive 2009/16/ec.

Erto, P., Lepore, A., Palumbo, B., Vitiello, L., 2015. A Procedure for Predicting and Controlling the Ship Fuel Consumption: Its Implementation and Test. Qual. Reliab. Eng. Int. 31. https://doi.org/10.1002/qre.1864

European Parliament and of the Council, 2012. on the monitoring and reporting of greenhouse gas emissions pursuant to Directive 2003/87/EC of the European Parliament and of the Counci, COMMISSION REGULATION (EU) .

IMO, 2009a. Guidance for the development of a ship energy efficiency management plan (SEEMP), MEPC.1/Circ.683.

IMO, 2009b. Guidelines for voluntary use of the ship Energy Efficiency Operational Indicator (EEOI), MEPC.1/Circ. 684.

IMO, 2000. Standard marine communication phrases (SMCP).

Inman, J., 1849. Navigation and nautical anstronomy for the use of the British seamen.

ITTC, 2017. Recommended Procedures and Guidelines: Preparation, Conduct and Analysis of Speed/Power Trials.

ITTC, 2008. Dictionary of Ship Hydrodynamics.

Kim, J.-G., Kim, H.-J., Jun, H.B., Kim, C.-M., 2016. Optimizing Ship Speed to Minimize Total Fuel Consumption with Multiple Time Windows. Math. Probl. Eng. 2016, 1–7. https://doi.org/10.1155/2016/3130291

Kim, J.-G., Kim, H.-J., Lee, P.T.-W., 2014. Optimizing ship speed to minimize fuel consumption. Transp. Lett. 6, 109–117. https://doi.org/10.1179/1942787514Y.0000000016

Lewis, E. V, 1988. Principles of Naval Architecture Second Revision - Volume II: Resistance, Propulsion and Vibration. The Society of Naval Architects and Marine Engineers.

Lu, R., Turan, O., Boulougouris, E., 2013. Voyage optimization, prediction of ship specific fuel consumption for energy efficient shipping, in: Low Carbon Shipping Conference. London, pp. 1–11.

MacGregor, J.F., Kourti, T., 1995. Statistical process control of multivariate processes. Control Eng. Pract. 3, 403–414. https://doi.org/10.1016/0967-0661(95)00014-L

Meng, Q., Du, Y., Wang, Y., 2016. Shipping log data based container ship fuel efficiency modeling. Transp. Res. Part B Methodol. 83, 207–229. https://doi.org/10.1016/J.TRB.2015.11.007

Montgomery, D.C., Peck, E.A., Vining, G.G., 2012. Introduction to linear regression analysis, Fifth Edit. ed. John Wiley & Sons, Hoboken, New Jersey.

Murphy, A.J., Weston, S.J., Young, R.J., 2012. Reducing fuel usage and CO2 emissions from tug boat fleets: Sea trials and theoretical modelling. Int. J. Marit. Eng. 154, A31–A41.

Nomikos, P., MacGregor, J.F., 1995a. Multivariate SPC Charts for Monitoring Batch Processes. Technometrics 37, 41–59. https://doi.org/doi:10.1016/0967-0661(95)00014-L

Nomikos, P., MacGregor, J.F., 1995b. Multi-way partial least squares in monitoring batch processes. Chemom. Intell. Lab. Syst. 30, 97–108. https://doi.org/10.1016/0169-7439(95)00043-7

Perera, L.P., Mo, B., 2016. Emission control based energy efficiency measures in ship operations. Appl. Ocean Res. 60, 29–46. https://doi.org/10.1016/J.APOR.2016.08.006

Petersen, J.P., Jacobsen, D.J., Winther, O., 2012. Statistical modelling for ship propulsion efficiency. J. Mar. Sci. Technol. 17, 30–39. https://doi.org/10.1007/s00773-011-0151-0

Smith, T.W.P., Parker, S., Rehmatulla, N., 2011. On the Speed of Ships, in: International Conference on Technologies, Operations, Logistics and Modelling for Low Carbon Shipping, LCS2011. University of Strathclyde, Glasgow, UK.

Trodden, D.G., Murphy, A.J., Pazouki, K., Sargeant, J., 2015. Fuel usage data analysis for efficient shipping operations. Ocean Eng. 110, 75–84. https://doi.org/10.1016/J.OCEANENG.2015.09.028

Veness, C., 2007. Calculating distance, bearing and more between two latitude/longitude points. Inst. Geophys. Texas.

Zaman, I., Pazouki, K., Younessi, S., Coleman, S., 2017. Development of automatic mode detection system by implementing the statistical analysis of ship data to

monitor the ... Int. J. Marit. Eng. 159, A225–A235. https://doi.org/10.3940/rina.ijme.2017.a3.411

FIGURES AND TABLES

	Variable	Description
Response variable	Y	Average fuel consumption per hour [Mt/h]
Predictor variables	V^3	SOG cubed [kn^3]
	σ_V^2	SOG variance [kn^2]
	W_H	Head wind [kn]
	W_F	Following wind [kn]
	W_S	Side wind [kn]
	SG_P	Shaft generator power (port) [kW]
	SG_S	Shaft generator power (starboard) [kW]
	ΔP	Power difference between two propeller shafts [kW]
	ΔSG	Power difference between two shaft generators [kW]
	T_{FD}	Departure draft (fore perpendicular) [m]
	T_{AD}	Departure draft (aft perpendicular) [m]
	T_{PD}	Departure draft (midship section – port) [m]
	T_{SD}	Departure draft (midship section – starboard) [m]
	T_{FA}	Arrival draft (fore perpendicular) [m]
	T_{AA}	Arrival draft (aft perpendicular) [m]
	T_{PA}	Arrival draft (midship section – port) [m]
	T_{SA}	Arrival draft (midship section – starboard) [m]
	$Trim_D$	Departure trim [m]
	$Trim_A$	Arrival trim [m]
	Δ	Displacement [t]
Other variables	UTC	Date and time UTC
	h	Actual navigation time [h]
	P^T	Shaft propeller power [kW]
	P^E	Electrical power [kW]
	M	Distance travelled [NM]
	h_{ij}	Running hours (i-th engine, j-th engine set) [h]
	$SFOC_{ij}$	Specific fuel oil consumption (i-th engine, j-th engine set)
	P_{ij}	Power (i-th engine, j-th engine set) [kW]
	η_j^m	Gearbox mechanical efficiency (j-th engine set)
	η_j^e	Shaft generator electrical efficiency (j-th engine set)

Table 1: Physical variables acquired at each voyage considered in the proposed approach.

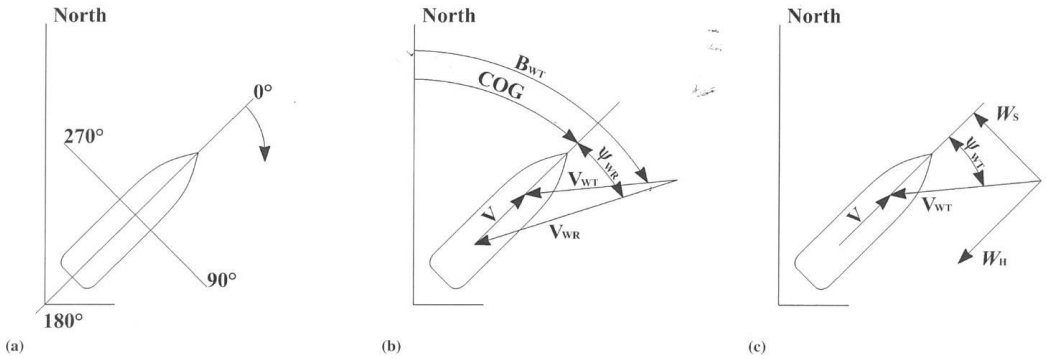

Figure 1: wind components

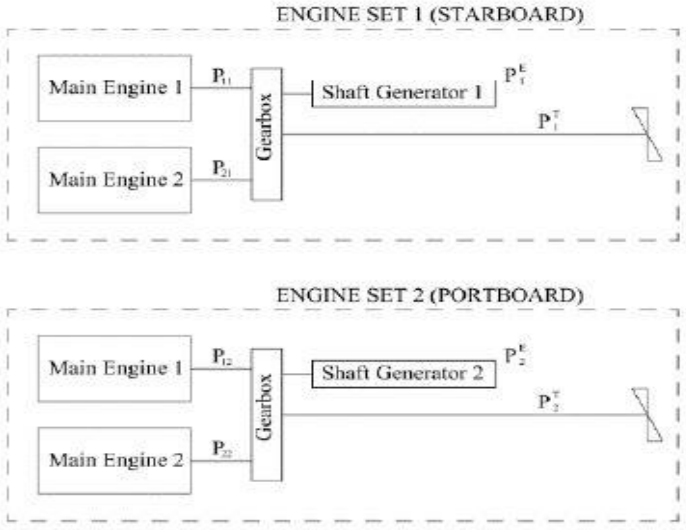

Figure 2: Engine room layout

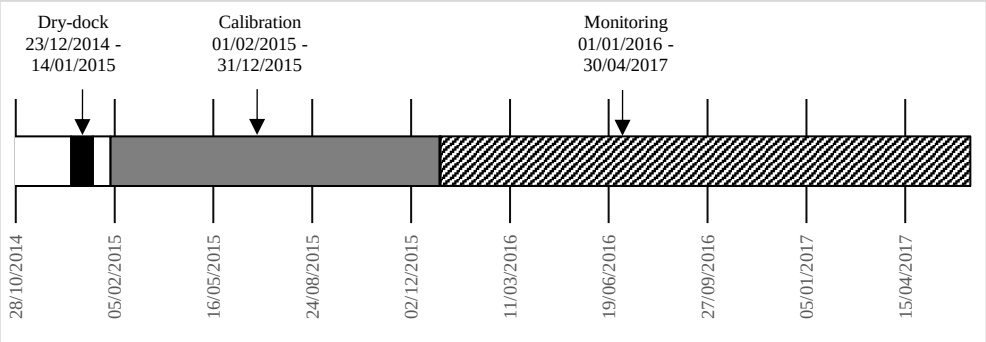

Figure 3: Ship 1 timeline

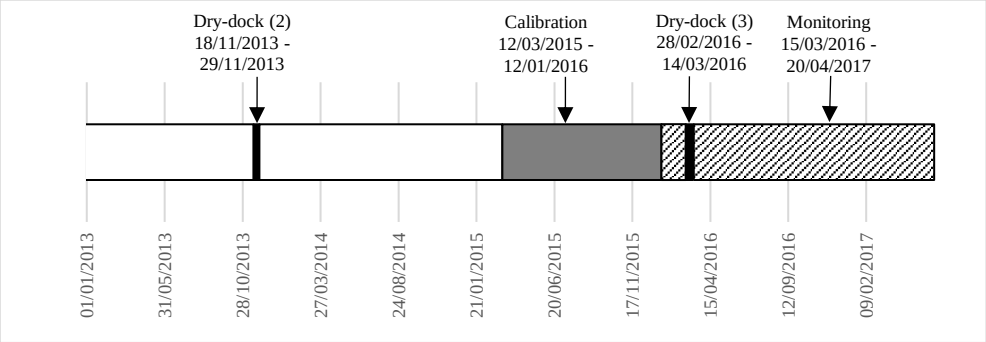

Figure 4: Ship 2 timeline

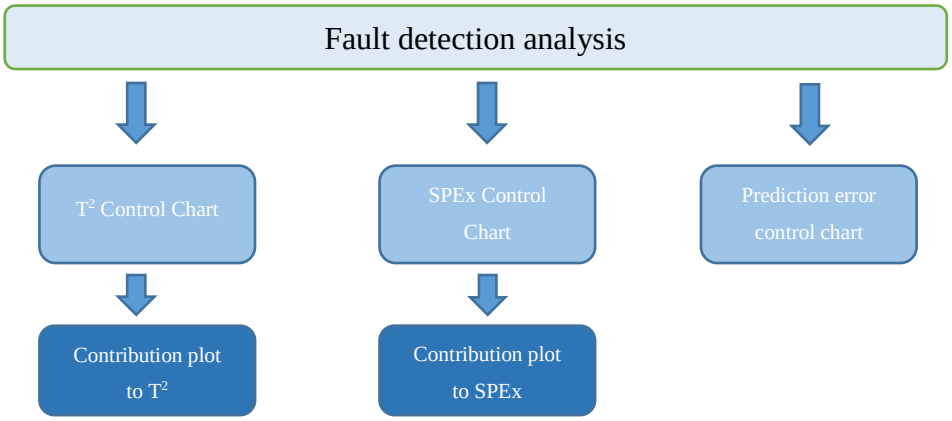

Figure 5: Fault detection analysis diagram

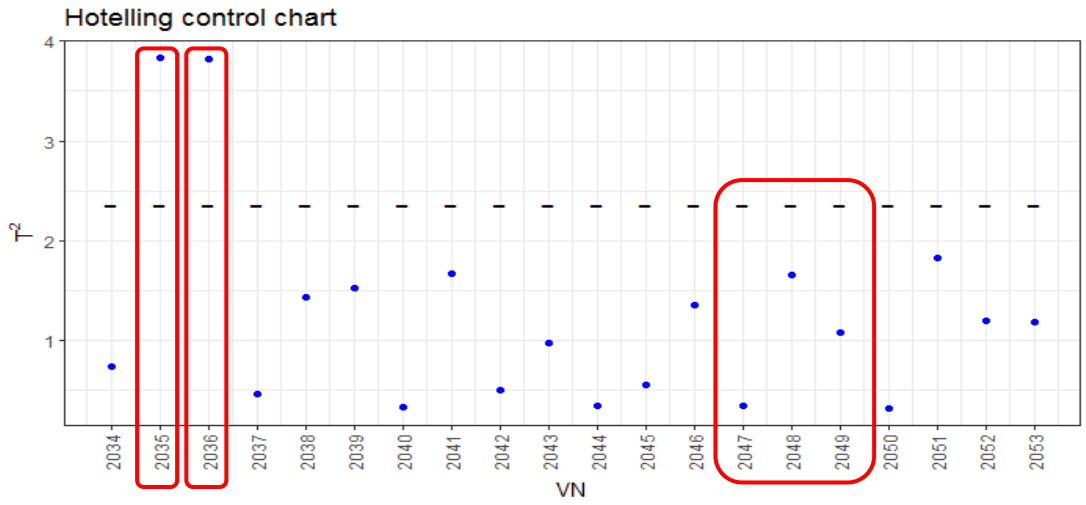

Figure 6: Hotelling T^2 control chart for voyages 2034—2053 of Ship 1.

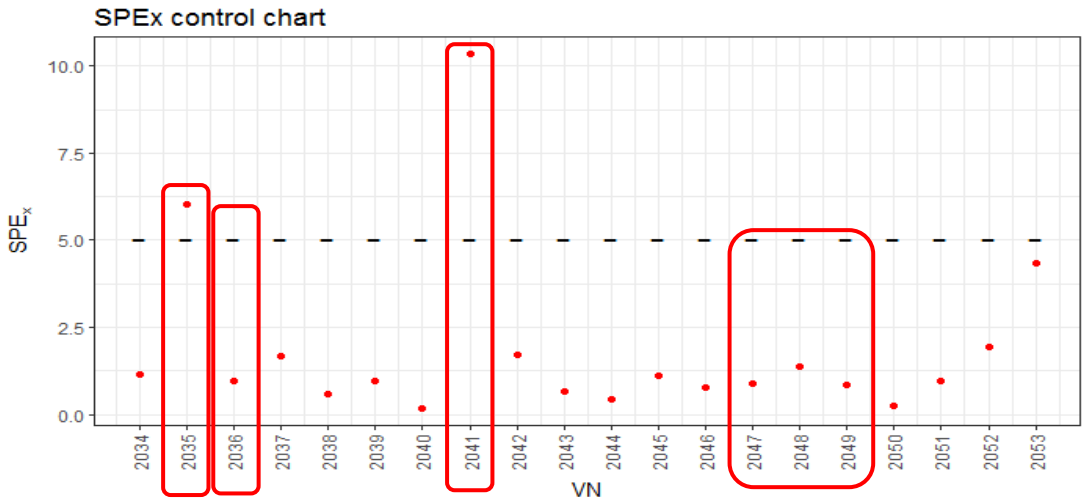

Figure 7: SPE_x control chart for voyages 2034—2053 of Ship 1.

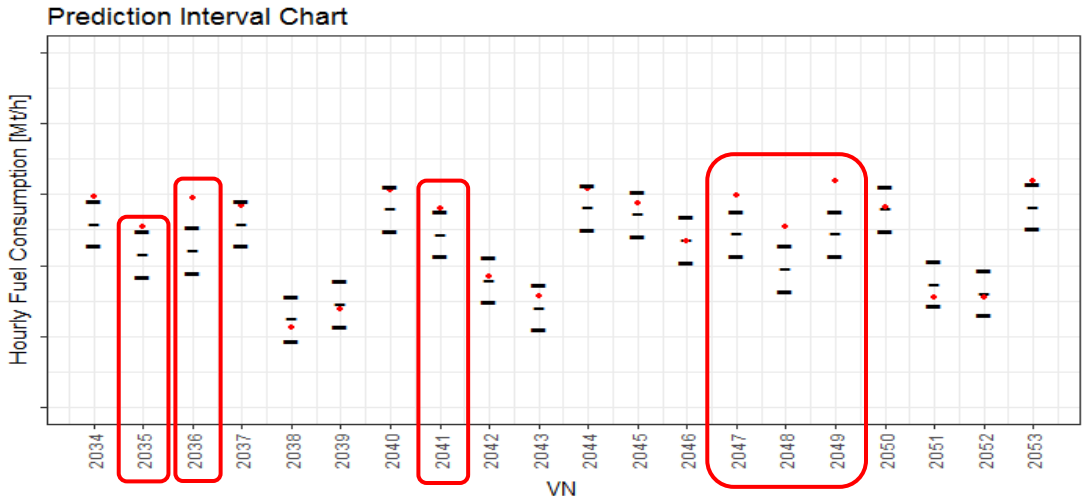

Figure 8: Prediction error control chart for voyages 2034—2053 of Ship 1.

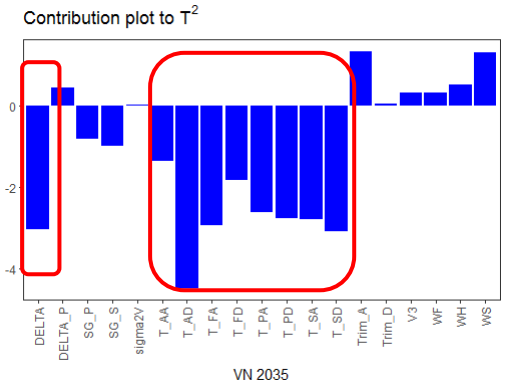

Figure 9: Contribution of the variables to the Hotelling T^2 statistic, for the voyage 2035 of Ship 1.

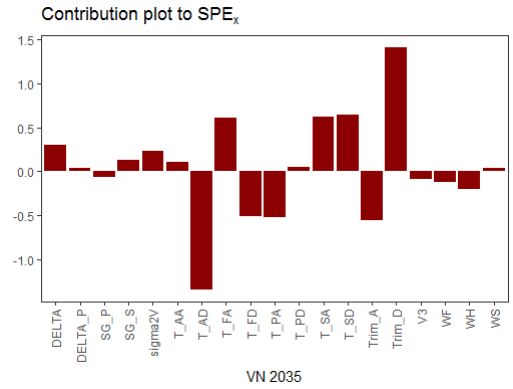

Figure 10: Contribution of the variables to the SPE_X statistic, for the voyage 2035 of Ship 1.

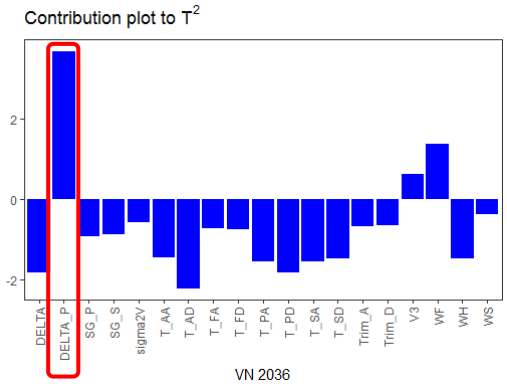

Figure 11: Contribution of the variables to the Hotelling T^2 statistic, for the voyage 2036 of Ship 1.

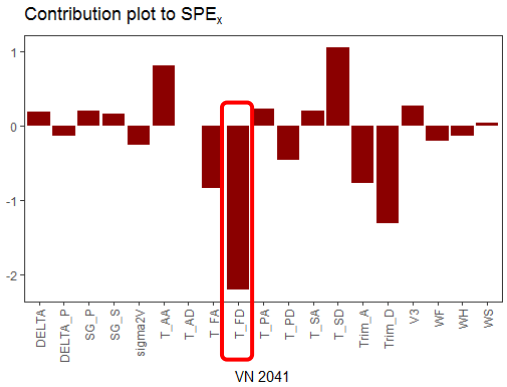

Figure 12: Contribution of the variables to the SPE_X statistic, for the voyage 2041 of Ship 1.

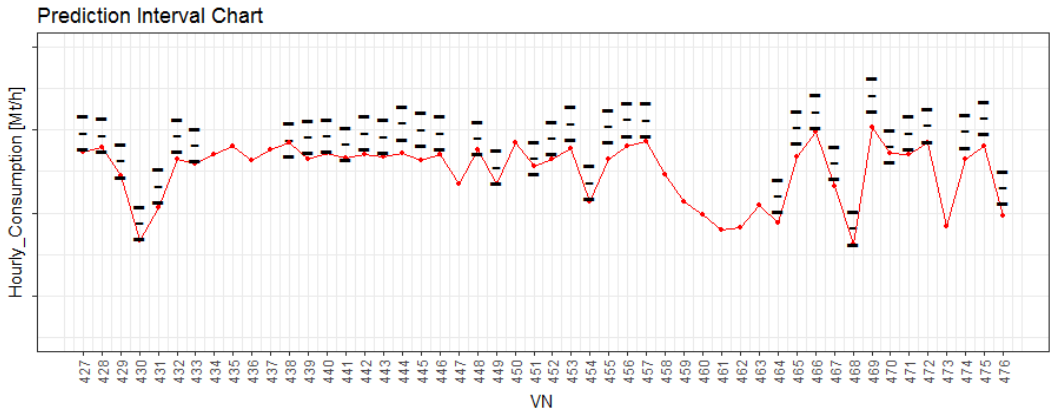

Figure 13: Prediction error control chart for voyages 427—476 of Ship 2, which follow an energy efficiency operation.