

Vessels Fuel Consumption: A Data Analytics Perspective to Sustainability

Andrea Coraddu, Luca Oneto, Francesco Baldi and Davide Anguita

Abstract The shipping industry is today increasingly concerned with challenges related with sustainability. CO₂ emissions from shipping, although they today contribute to less than 3% of the total anthropogenic emissions, are expected to rise in the future as a consequence of increased cargo volumes. On the other hand, for the 2 °C climate goal to be achieved, emissions from shipping will be required to be reduced by as much as 80% by 2050. The power required to propel the ship through the water depends, among other parameters, on the trim of the vessel, i.e. on the difference between the ship's draft in the fore and the aft of the ship. The optimisation of the trim can, therefore, lead to a reduction of the ship's fuel consumption. Today, however, the trim is generally set to a fixed value depending on whether the ship is sailed in loaded or ballast conditions, based on results performed on model tests in basins. Nevertheless, the on-board monitoring systems, which produce a huge amount of historical data about the life of the vessels, lead to the application of state of the art data analytics techniques. The latter can be used to reduce the vessel consumption by means of optimising the vessel operational conditions. In this book chapter, we present the potential of data-driven based techniques for accurately predicting the influence of independent variables measured from the on board monitoring system and the fuel consumption of a specific case study vessel. In particular, we show that gray-box models (GBM) are able to combine the high prediction accuracy of black-box models (BBM) while reducing the amount of data required for training the

A. Coraddu (✉)

R&D Department, DAMEN Shipyard Singapore, Singapore, Singapore
e-mail: andrea.coraddu@damen.com

L. Oneto · D. Anguita

DIBRIS—University of Genoa, Genoa, Italy
e-mail: luca.oneto@unige.it

F. Baldi

Department of Shipping and Marine Technology, Chalmers University
of Technology, Gothenburg, Sweden
e-mail: francesco.baldi@chalmers.se

D. Anguita

e-mail: davide.anguita@unige.it

© Springer International Publishing AG 2018

C. Cruz Corona (ed.), *Soft Computing for Sustainability Science*,

Studies in Fuzziness and Soft Computing 358, DOI 10.1007/978-3-319-62359-7_2

model by adding a white-box model (WBM) component. The resulting GBM model is then used for optimising the trim of the vessel, suggesting that between 0.5 and 2.3% fuel savings can be obtained by appropriately trimming the ship, depending on the extent of the range for varying the trim.

1 Introduction

Shipping is a relatively efficient mean of transport when compared to other transport modes [75]. Despite its efficiency, however, shipping contributes significantly to air pollution [13], mainly in the form of sulphur oxides, nitrogen oxides, particulate matter, and carbon dioxide. For the latter, the contribution from shipping to global emissions is required to decrease significantly in the coming years [6, 20]. Since greenhouse gas emissions from the combustion of oil-based fuels are directly proportional to fuel consumption, improving ship energy efficiency is one of the possible solutions to this issue. Measures for the improvement of ship energy efficiency are normally divided into design and operational measures. While the former have been associated to larger saving potential, the latter can still provide a significant reduction in fuel consumption, while requiring a much more limited capital investment [6]. However, the large amount of variables influencing ship energy efficiency makes it hard to assess ship performance in relation to a standard baseline. Operational measures include, among others, improvement in voyage execution, reduction of auxiliary power consumption, weather routing, optimised hull and propeller polishing schedule, slow steaming, and trim optimisation [2, 42, 64].

Among the above mentioned fuel saving measures, trim optimisation has been extensively discussed in the past. It is well known, from hydrodynamics principles, that the trim of the vessel can significantly influence its fuel consumption [60]. In most cases principles of trim optimisation are applied roughly; the crew is provided with an indicative value for the trim to use when sailing laden and when sailing ballast, based on model tests. However, many more factors can influence the optimal value of the trim, such as draught, weather conditions, speed [38]. Taking these aspects into account when selecting the appropriate trim can therefore lead to significant, cost-free savings in terms of fuel required for ship propulsion. Previous work in scientific literature related to trim optimisation have focused on two main alternative strategies: White-Box numerical Models (WBMs), and black-box numerical models (BBMs). WBMs describe the behaviour of the ship resistance, propeller characteristics and engine performances based on governing physical laws and taking into account their mutual interactions [49]. The higher the detail in the modelling of the physical equations which describe the different phenomena, the higher the expected accuracy of the results and the computational time required for the simulation. WBMs are generally rather tolerant to extrapolation and do not require extensive amount of operational measurements; on the other hand, when employing models that are computationally fast enough to be used for online optimisation, the expected accuracy in the prediction of operational variables is relatively low. In addition, the construction of the

model is a process that requires competence in the field, and availability of technical details which are often not easy to get access to. Examples of the use of WBMs for the optimisation of ship trim are [46], who employed advanced Computational Fluid Dynamics (CFD) methods, and [54] who employed simpler empirical models (Holtrop-Mannen) for the estimation of possible gains from trim optimisation. Differently from WBMs, BBMs (also known as data driven models [74]), make use of statistical inference procedures based on historical data collection. These methods do not require any a-priory knowledge of the physical system and allow exploiting even measurements whose role might be important for the calculation of the predicted variables but might not be captured by simple physical models. On the other hand, the model resulting from a black-box approach is not supported by any physical interpretation [65], and a significant amount of data (both in terms of number of different measured variables and of length of the time series) are required for building reliable models [12]. As an example, in [58] an application of BBMs is proposed (in particular of artificial neural network) to the prediction of the fuel consumption of a ferry and applied to the problem of trim optimisation. Gray-box models (GBMs) have been proposed as a way to combine the advantage of WBMs and BBMs [11]. According to the GBMs principles, an existing WBM is improved using data-driven techniques, either in order to calculate uncertain parameters or by adding a black-box component to the model output [48]. GBMs allow exploiting both the mechanistic knowledge of the underlying physical principles and available measurements. The proposed models are more accurate than WBMs with similar computational time requirements, and require a smaller amount of historical data when compared to a pure BBMs.

The aim of this book chapter is to propose the application of a gray-box modelling approach to the prediction of ship fuel consumption which can be used as a tool for online trim optimisation. In this framework the authors exploit Machine Learning techniques based on kernel methods and ensemble techniques [5, 65] so to improve an effective but simplified physical model [8] of the propulsion plant. The proposed model is tested on real data [3] collected from a vessel during two years of on board sensors data acquisitions (e.g. sheep speed, axis rotational speed, torque, wind intensity and direction, temperatures, pressure, etc.).

2 The Sustainability Challenge in Shipping

2.1 The Shipping Sector

International trade has been a major factor in the development of mankind all throughout the history. This can be seen in particular today, as shipping contributes to approximately 80–90% of global trade (in ton km, [52]) with an increase from 2.6 to 9.8 billion tons of cargo from 1970 to 2014. Today anything from coal, iron ore, oil

and gas to grains, cars, and containerized cargo is transported by sea, thus making shipping the heart of global economy [73].

If compared to other transportation modes, shipping is relatively efficient if measured in terms of fuel consumed per unit of cargo transported and of distance covered [75]. Nevertheless, shipping is today under strong pressure for reducing its fuel consumption, both from an environmental and economical perspective.

2.2 Shipping and Carbon Dioxide Emissions

The main connection between energy efficiency and sustainability in shipping relates to the emissions of greenhouse gas (GHG), that are considered today to be the main contributor to global warming. Despite carbon dioxide (CO₂) emissions from the shipping sector were estimated to amount to less than 3% of the total in 2012, they are expected to grow in the future by between 50 and 250% in relation to the expected increase of transport volumes [67].

This pressure related to making shipping more sustainable will also have more and more impact on the shipping industry from an economical perspective. Not only environmental regulations are becoming stricter in many areas of the world (compliance often requires higher fuel expenses). In the particular case of CO₂ emissions, market based measures are being discussed, particularly but not only in the European Union, as a mean for incentivising the transition to low-carbon shipping.

In relation to the strive for sustainability, shipping is a very peculiar business, where conditions are not optimal for incentivising energy efficiency. Split incentives are often a hinder to implementing energy efficiency measures, as neither the owner of the ship, nor its operator pay for the fuel [39]. In addition, differently from e.g. planes and cars, ships are built on individual or small-series basis; this makes it particularly expensive to invest into research and development on an individual ship basis [18, 76]. Ships are very long-lasting products, whose operational life can range from 15 to more than 30 years [69].

In these conditions, although technical improvements to ship energy systems (both by retrofitting and in the design phase) are seen as the solutions with the largest potential for reducing ship fuel consumption, operational measures are of particular interest. In fact they do not require any initial investment and, therefore, are particularly easy to implement [6].

2.3 Operational Efficiency in Shipping

Operational measures is a category that includes different measures for energy efficiency on board that do not require the installation of new equipment. On most vessels, the energy demand for propulsion represents the largest share of the total energy demand [3]. For this reason, most of the measures that aim at reducing ship

fuel consumption relate to the reduction of the fuel demand from the main engines. This, in turn, can be achieved by either reducing the thrust required for moving the ship's hull through the water, or by improving the efficiency of the most relevant conversion components, namely the propeller and the engine. An appropriate optimisation requires, however, an in-depth understanding of the influence of the speed of the vessel on its fuel consumption in different environmental and operational conditions.

As the power demand for propulsion roughly depends on the ship's speed to the third power (up to the fourth power for faster ships), reducing the speed of the vessel is often regarded as a possible solution for improving energy efficiency. Although the practice of slow-steaming has its inconveniences (e.g. demand for more ships to be built, longer time at sea, higher inventory costs), it has been shown that fuel can be saved by optimising the speed at each instant of the voyage, without changing the total voyage time. This practice is normally referred to as weather routing. Its correct application requires, however, not only the availability of reliable short-middle term predictions of the weather conditions, but also of an accurate understanding of the influence of given weather conditions on the ship's power demand for propulsion.

For given conditions of ship speed and weather, there are other operational parameters that influence the power demand for propulsion. In particular, the trim (defined as the difference between the draft at the ship's fore and aft, thereby measuring how much the position of the ship differentiates from that of being parallel to the sea surface) can be optimised in order to adapt to conditions of minimal demand for propulsive power. This is normally done on board starting from rules of thumb based on tests performed on ship physical models at reduced scale, where the ship's average draft and speed are the parameters that most influence the choice of the optimal trim. However, in real operating conditions, not only the influence of these variables can be different from what predicted by model tests, but also other conditions (e.g. the weather) can play a role in the determination of the optimal draft.

The added resistance coming from the growth of different types of organisms on the surface of the hull also plays a major role on the total power demand for ship propulsion, which can increase by up to 100% [63] as a consequence of the increased hull roughness. As a solution to this issue, most ships use a thin layer of poisonous paint (normally referred to as antifouling paint) which slowly releases substances which are poisonous for the organisms that grow on the surface of the hull. In addition, the hull is cleaned and, if necessary, re-painted at specific intervals. The choice of the hull cleaning intervals is, today, mostly based on rules of thumb (e.g. once a year), generally as a consequence of the difficulty of predicting the relative contribution of fouling on the total ship resistance. This practice could therefore be substantially improved if the contribution of added resistance due to fouling to the total ship resistance could be evaluated more accurately.

Finally, a significant share of ships are today equipped with a controllable pitch propeller (CPP), i.e. a propeller where the inclination of the blades in relation to the propeller axis can be changed according to the specific requirements as an operational variables. When this type of propeller is installed, the choice of the pitch is normally pre-set as a function of the propeller speed for optimising the efficiency of

the propeller. However, not only the optimal propeller efficiency is also influenced by other factors (e.g. ship draft and weather conditions), but also the engine efficiency is influenced by the choice of its operating conditions in terms of speed and torque requirements. This shows potential for additional fuel savings if the propeller pitch is continuously optimised for optimal efficiency of the entire propulsion train.

It appears clearly from the previous section that an appropriate ability of predicting the influence of all the different environmental and operational variables on the performance of the ship is of utmost importance for achieving the most out of different operational measures for improving ship energy efficiency.

3 Problem Description

3.1 Ship Description

In this book chapter, the authors propose the utilisation of a predictive model of the fuel consumption for the online optimisation of the trim of a vessel. The proposed method has been tested on a Handymax chemical/product tanker in order to show its potential. A conceptual representation of the ship propulsion plant is shown in Fig. 1, while relevant ship features are presented in Table 1. The ship systems consists of two main engines (MaK 8M32C four-stroke Diesel engines) rated 3840 kW each and designed for operation at 600rpm. The two engines are connected to one common gearbox; the gearbox has two outputs: a controllable pitch propeller designed for operations at 105 rpm for ship propulsion; and a shaft generator (rated 3200 kW)

Fig. 1 Conceptual representation of the ship propulsion system

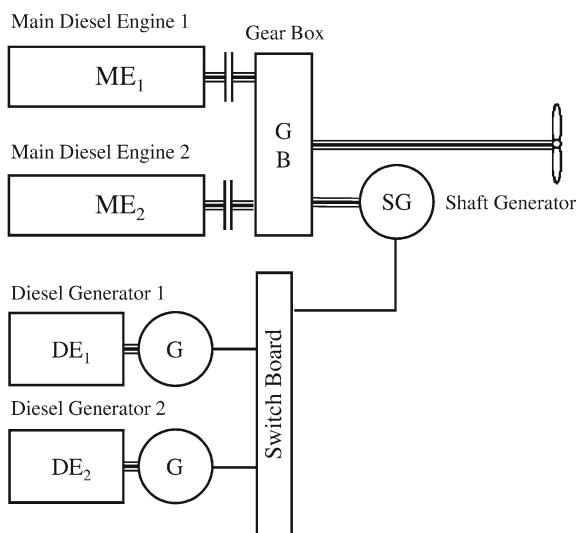
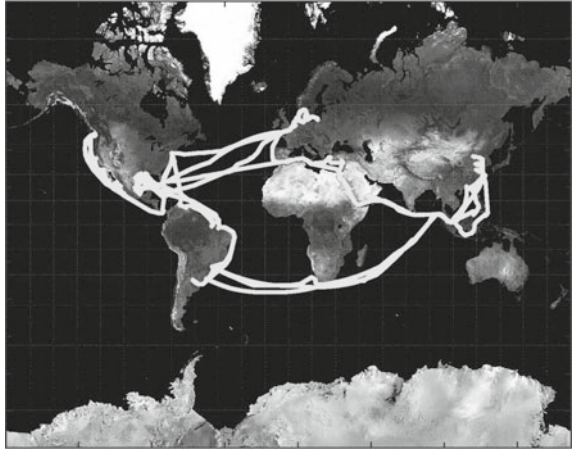


Table 1 Main features of the case study ship

Ship feature	Value	Unit
Deadweight	47000	(t)
Installed power (Main Engines)	3840 (x2)	[kW]
Installed power (Auxiliary Engines)	682 (x2)	[kW]
Shaft generator design power	3200	[kW]
Exhaust boilers design steam gen.	1400	[kg/h]
Auxiliary boilers design steam gen.	28000	[kg/h]

Fig. 2 Description of the ships routes

used for fulfilling on board auxiliary power demand. Auxiliary power can also be generated by two auxiliary engines rated 682 kW each. Auxiliary heat demand is fulfilled by a combination of exhaust gas boilers and auxiliary oil-fired boilers.

The ship is mainly used in the spot market (i.e. based on short-term planning of ship logistics, as opposed to long-term agreements with cargo owners on fixed schedules and routes) and therefore operates according to a variable schedule, both in terms of time spent at sea and of ports visited. The variety of different routes is shown in Fig. 2. Figures 3 and 4 represent the observed ship operations for the selected time period. It can be seen that although the ship spends a significant part of time in port, most of ship operations are related to open sea transport, either in laden or ballast mode (see Fig. 3). The focus of this work lies in the optimisation of ship trim; consequently, only data points related to sailing operations are considered in this study. Operations of manoeuvring, cargo loading, cargo unloading, and port stays were therefore excluded from the original dataset. These transport phases happen at a broad range of speeds, as shown in Fig. 4, which provides additional evidence of the need for an efficient tool for the optimisation of ship operations in different operational conditions. Because of their specific trading pattern, tankers are normally used in two very distinct operational modes: laden (i.e. with full cargo holds, delivering liquid bulk cargo to the destination port), and ballast (with empty

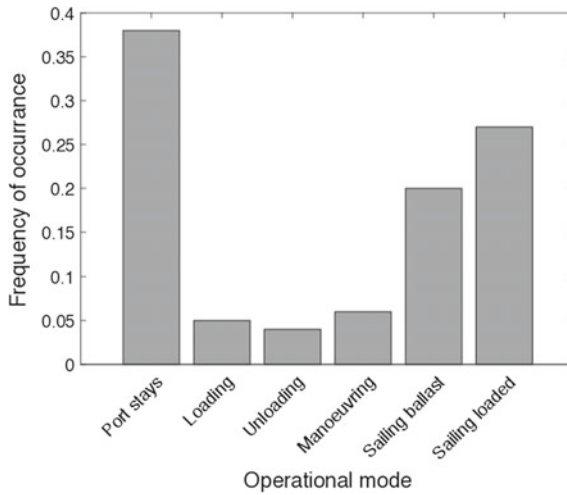


Fig. 3 Time spent in each operational mode for the selected vessel in the chosen period

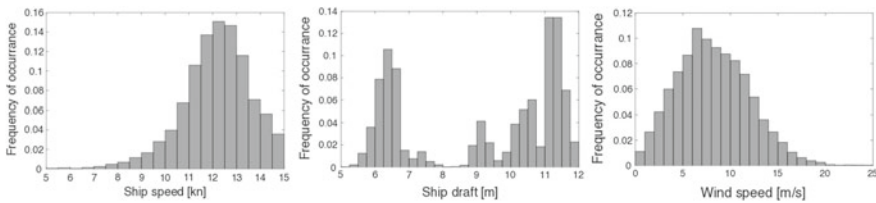


Fig. 4 Speed, draft and wind distributions during sailing time for the selected vessel in the chosen period

cargo holds, sailing to the port where the next cargo is available for loading). In reality, even when loaded, tankers vessels do not always sail with completely full holds due to differences in order sizes. The ship’s draught can consequently vary, depending on the operation, from 11 m when the ship is fully loaded to 6 m when cargo holds are completely empty. The distribution of ship draught over the proposed dataset is presented in Fig.4. In addition to ship speed and draught, weather conditions are also known to have an influence on the optimal trim to be used when sailing, and can vary during ship operations. Figure 4 represents wind speed which, in turn, is strongly correlated to ship added resistance.

3.2 Data Logging System

The ship under study is provided with a data logging system installed by an energy management provider which is used by the company both for on board monitoring and for land-based performance control. Table 2 summarises the available measurements from the continuous monitoring system.

The original data frequency measured by the monitoring system is of 1 point every 15 s. In order to provide easier data handling, the raw data are sent to the provider server, where they are processed into 15 min averages. The data processing is performed by the provider company and could not be influenced or modified by the authors.

Measured values come from on board sensors, whose accuracy and reliability cannot be ensured in the process. In particular, issues related to the measurement of speed through water (LOG speed) are well known. Such measurements are often partly unreliable since the flow through the measurement device can be easily disturbed by its interaction with the hull or by other environmental conditions. On the other hand measurements of speed over ground (GPS speed), although more reliable, do not include the influence of currents, which can be as strong as 2÷3 knots depending on time and location and therefore influence ship power demand for propulsion. Fuel consumption is measured using a mass flow meter, which is known to be more accurate of the more common volume flow meters as it eliminates uncertainty on fuel density. It should be noted, however, that measurements of fuel specific energy content (LHV) were not available; variation of heavy fuel oil LHV is known to be in the order of ± 2 MJ/kg, which corresponds to a variation of $\pm 5\%$. Propeller speed, torque measurement and fuel mass flow accuracy were provided by the shipyard at respectively ± 0.1 , ± 1 and $\pm 3\%$.

4 From Inference to Data Analytics

Inference is the act or process of deriving logical conclusions from premises known or assumed to be true [51]. There are two main families of inference processes: deterministic, and statistical inference. The former studies the laws of valid inference, while the latter allows to draw conclusions in the presence of uncertainty, and therefore represents a generalisation of the former. Several different types of inference are commonly used when dealing with the conceptual representation of reality as shown in Fig. 5:

- Modelling/approximation refers to the process of building a model of a real system based on the knowledge of the underlying laws of physics that are known to govern the behavior of the system. Depending on the expected use and needs of the model, as well as on the available information, different levels of approximation can be used. Modelling/approximation of a real system only based on mechanistic knowledge can be categorised as deterministic inference [37];

Table 2 Measured values available from the continuous monitoring system

Variable name	Unit
Time stamp	
Latitude	(°)
Longitude	(°)
Fuel consumption (Main engines)	(kg/15 mins)
Auxiliary engines power output	(kW)
Shaft generator power	(kW)
Propeller shaft power	(kW)
Propeller speed	(kW)
Ship draft (fore)	(m)
Ship draft (aft)	(m)
Draft port	(m)
Draft starboard	(m)
Relative wind speed	(m/s)
Relative wind direction	(°)
GPS heading	(°)
Speed over ground (GPS)	(kn)
Speed through water (LOG)	(kn)
Sea depth	(m)
Sea water temperature	(°C)
CPP setpoint	(°)
CPP feedback	(°)
Fuel density	(kg/m ³)
Fuel temperature	(°)
Ambient pressure	(mbar)
Humidity	(%)
Dew point temperature	(°C)
Shaft torque	(kNm)
Rudder angle	(°)
Acceleration X direction	(-)
Acceleration Y direction	(-)
Acceleration Z direction	(-)
GyroX	(-)
GyroY	(-)
GyroZ	(-)
Roll	(-)
Pitch	(-)
Yaw	(-)

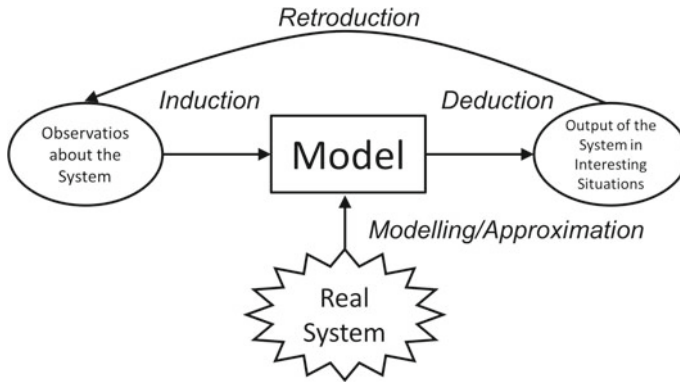


Fig. 5 Type of inference exploited in this book chapter

- When the model is built by statistically elaborating observations of system inputs and outputs, the process belongs to the category of statistical induction. As the model is inferred based on measurements affected by different types of noise, this process is intrinsically under the effect of uncertainty and therefore belongs to the category of statistical inference [74];
- The process of using an existing model to make predictions about the output of the system given a certain input is called deductive inference. This process can be both deterministic or probabilistic depending on how the model is formulated [14];
- The process of actively modifying model inputs in order to obtain a desired output is normally referred to as retrodution (or abduction) [41].

The subject of this book chapter can hence be seen as the application of a general category of problems to a specific case. The physical laws governing ship propulsion are known and widely used in the dedicated literature with the purpose of modelling the ship behaviour [11]. Moreover a series of historical data about the ship's propulsion system are available, and based on this it is possible to build a statistical model of the process [21, 32, 45, 68] which again can be exploited to predict the behaviour of the system. In particular data analytics tools allow performing different levels of statistical modelling [21]:

- descriptive analytics tools allow understanding what happened to the system (e.g. what was the temperature of the cylinders of the engine in the last days). Descriptive analytics answers to the question 'What happened?'
- diagnostic analytics tools allow understanding why something happened to the system (e.g. the fuel consumption it too high and this is due to a the decay of the hull). Diagnostic analytics answers to the question 'Why did it happen?'
- predictive analytics tools allow making predictions about the system (e.g. when a new propeller is installed to reduce fuel consumption). Predictive analytics answers to the question 'What will happen?'

- prescriptive analytics tools allow understanding why the system behave in a particular way and how to force the system to be in a particular state (e.g. what is the best possible way to steer the ship in order to save fuel). Prescriptive analytics answers to the question ‘How can we make it happen?’

Descriptive analytics is something very simple to implement, for example in Sect. 3 authors showed some compressed information coming from the historical data collection which can be interpreted as a descriptive analytic process. These tools are the least interesting ones since there is no additional knowledge extracted from the data [21]. Diagnostic analytics is a step forward where the authors try to understand what happened in the past, searching correlation in the data in order to get additional information from the data itself. Examples of these approach in the context of naval transportation system can be found in [12, 40, 57, 78]. Finally, predictive and prescriptive analytics are the most complex approaches where a model of the system is built and studied in order to understand the accuracy and the properties of the model and make the system behave in a particular way. This is the most important analysis in practical applications since, even if diagnostic analytics allows improving the understanding of past and present conditions of the system, it is more important to predict the future and take action in order to prevent the occurrence of some event [12, 24] (substitute a component before it fails) or to make some event happen [58] (reduce the fuel consumption of a ship).

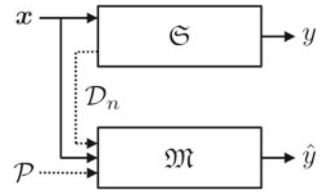
For these reasons in the next sections a more rigorous framework is depicted together with the description of the approaches adopted for building predictive models. An assessment of their accuracy and properties is performed and a complete description about how to use these models to force the system in producing an output is provided.

4.1 Supervised Learning

In the context of supervised machine learning, we are interested in a particular sub-problem which is the regression one. Regression helps to understand how the value of a dependent variable changes when any one of the independent variables is varied. Using the conventional regression framework [65, 74] a set of data $\mathcal{D}_n = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$, with $\mathbf{x}_i \in \mathcal{X} \subseteq \mathbb{R}^d$ and $y_i \in \mathcal{Y} \subseteq \mathbb{R}$, are available from the automation system. Each tuple (\mathbf{x}_i, y_i) is called sample and each element of the vector $\mathbf{x} \in \mathcal{X}$ is called feature.

When inferring a model starting from a real system, the goal is to provide an approximation $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{Y}$ of the unknown true model $\mathcal{S} : \mathcal{X} \rightarrow \mathcal{Y}$. \mathcal{S} and \mathcal{M} are graphically represented in Fig. 6. It should be noted that the unknown model \mathcal{S} can be also seen, from a probabilistic point of view, as a conditional probability $\mathbb{P}(y|\mathbf{x})$ or, in other words, as the probability of the output y given the fact that we observed \mathbf{x} as an input to \mathcal{S} .

Fig. 6 The regression problem



As previously described, in this book chapter three alternative modelling strategies are compared: white-, black-, and gray-box models:

- **White Box Model (WBM):** in this case the model $\mathfrak{M}_{\text{WBM}}$ is built based on a priori, mechanistic knowledge of \mathfrak{S} (numerical description of the body hull, propulsion plant configuration, design information of the ship). The implementation of a WBM in this specific case is described in Sect. 5.1.
- **Black Box Model (BBM):** in this case the model $\mathfrak{M}_{\text{BBM}}$ is built based on a series of historical observation of \mathfrak{S} (or in other words \mathcal{D}_n). In this book chapter, this is done by exploiting state of the art Machine Learning techniques as described in Sect. 5.2.
- **Gray Box Model (GBM):** in this case the WBM and BBM are combined in order to build a model $\mathfrak{M}_{\text{GBM}}$ that takes into account both a priori information and historical data \mathcal{D}_n so to improve the performances of both the WBM and BBM models. The implementation of the GBM principle to the specific case of this work is described in Sect. 5.3.

4.2 Estimation of Model Accuracy

The accuracy of the model \mathfrak{M} as a representation of the unknown system \mathfrak{S} can be evaluated using different measures of accuracy [26]. In particular, given a series data $\mathcal{T}_m = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\}$,¹ the model will predict a series of outputs $\{\hat{y}_1, \dots, \hat{y}_m\}$ given the inputs $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$. Based on these outputs it is possible to compute these performance indicators:

- mean absolute error (MAE) $\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i|$
- mean absolute percentage error (MAPE) $\text{MAPE} = 100 \frac{1}{m} \sum_{i=1}^m \left| \frac{y_i - \hat{y}_i}{y_i} \right|$
- mean square error (MSE) $\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$
- normalised mean square error (NMSE) $\text{NMSE} = \frac{1}{m\Delta} \sum_{i=1}^m (y_i - \hat{y}_i)^2$,
 $\Delta = \frac{1}{m} \sum_{i=1}^m (y_i - \bar{y})^2$, and $\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i$

¹The set \mathcal{T}_m must be a different set respect to \mathcal{D}_n which has been used to built the model \mathfrak{M} in the case of BBMs and GBMs [1].

- relative error percentage (REP) $REP = 100 \sqrt{\frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m y_i^2}}$
- Pearson product-moment correlation coefficient (PPMCC) which allows to compute the correlation between the output of the system and the output of the model $PPMCC = \frac{\sum_{i=1}^m (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^m (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^m (\hat{y}_i - \bar{\hat{y}})^2}}$, and $\bar{\hat{y}} = \frac{1}{m} \sum_{i=1}^m \hat{y}_i$

Note that all these measures of accuracy are useful for giving an exhaustive description of the quality of the forecast [26].

4.3 Prescriptive Analytics

Once the model \mathfrak{M} of the system \mathfrak{S} is available, it is possible to control its inputs in order to produce a desired output. In this particular application, the goal is to find the minimum for the fuel consumption by acting on the ship's trim while keeping all other model inputs unchanged.

This approach, however, requires additional care and understanding of the underlying physics of the system:

- With reference to the previous work from the authors [11], not all variables available as measurements can be used as predictors. In this case, in particular, the power and torque at the propeller shaft had to be excluded from the input list (see Table 3). Changing the trim would consequently change ship resistance and, therefore, the power required for its propulsion. Therefore modifying the trim while keeping the propeller power constant would represent a conceptual error.
- Not all possible trim values are physically allowed, and therefore boundary values, based on a priori knowledge of the system, should be provided.
- Although GBMs are more reliable in the extrapolation phase, their accuracy is expected to be reduced if they are extrapolated too far for outside the boundaries of the original range \mathcal{D}_n . Extrapolation is therefore allowed (the use of GBMs proposed in this book chapter is also based on their improved performance for extrapolation compared to BBMs) but this operation should be performed with care.

Based on these considerations, in this book chapter a method for trim optimisation is proposed. WBM, BBM and GBM are presented and compared based on the accuracy metrics proposed in Sect. 4.2. Based on this comparison, one model is selected for further analysis, checked for physical plausibility (Sect. 6) and used for application to the problem of trim optimisation (Sect. 7).

Table 3 Variable of Table 2 exploited to built the \mathfrak{M}

Id	Name	Type
1	Latitude	Input
2	Longitude	Input
3	Volume	Input
4	State	Input
5	Auxiliary consumed	Input
6	Auxiliary electrical power output	Input
8	Shaft rpm	Input
9	Ship draft (fore)	Input
10	Ship draft (aft)	Input
11	Relative wind speed	Input
12	Relative wind direction	Input
13	GPS heading	Input
14	GPS speed	Input
15	Log speed	Input
16	Shaft generator power	Input
17	Sea depth	Input
18	Draft Port	Input
19	Draft Starboard	Input
20	Sea water temperature	Input
21	CPP setpoint	Input
22	CPP feedback	Input
23	Fuel density	Input
24	Fuel temperature	Input
25	Ambient pressure	Input
26	Humidity	Input
27	Dew point temperature	Input
29	Rudder angle	Input
30	Acceleration X direction	Input
32	Acceleration Y direction	Input
32	Acceleration Z direction	Input
33	GyroX	Input
34	GyroY	Input
35	GyroZ	Input
36	Roll	Input
37	Pitch	Input
38	Yaw	Input
39	True direction	Input
40	True speed	Input

(continued)

Table 3 (continued)

Id	Name	Type
41	Beaufort	Input
	Shaft power	Output
	Shaft torque	Output
	Main engine consumption	Output
	Shaft power predicted by the WBM	Input GBMs
	Shaft torque predicted by the WBM	Input GBMs
	Main engine consumption predicted by the WBM	Input GBMs

5 White, Black and Gray Box Models

5.1 White Box Models

A numerical model, the so called White Box Model (WBM), based on the knowledge of the physical processes was developed by the authors. The WBM model is able to evaluate the ship consumption, for different ship speed V and displacement Δ in calm water scenario.

The model is based on the knowledge of the ship's hull geometry, mass distribution, propeller characteristics and main Diesel engine consumption map. The selected control variables (i.e. the system input which is under the user's control) taken into account are: the main engine revolution N and the pitch ratio P/D . The control of these variables allow the ship to sail at the desired speed. The total ship's fuel consumption is used as model output.

The core of the procedure is the engine-propeller matching code utilised to evaluate the total ship fuel consumption and already tested as an effective tool in a previous work [9].

The prediction of ship resistance in calm water can be performed according to different approaches, normally divided in parametric approaches [4, 29, 34, 35], and approaches based on computational fluid dynamics (CFD), such as the Reynolds averaged Navier-Stokes (RANS) or boundary element methods (BEM) [33]. In this study only parametric methods were considered because of their lower computational requirements. In particular, the Guldhammer Harvald method [29] was employed for the prediction of calm water resistance and, in particular, of the coefficient of total hull resistance in calm water C_T in Eq. (1). The inputs related to ship geometry used in the Guldhammer Harvald method are summarised in Table 4.

$$R_{\text{tot}} = \frac{1}{2} C_T \rho S V^2 \quad (1)$$

where ρ is the sea water density.

Table 4 Main input quantities for ship resistance prediction

Input	Symbol	Unit
Length on waterline	L_{WL}	(m)
Breadth on waterline	B_{WL}	(m)
Draught	T	(m)
Volume	V	(m ³)
Wetted surface	S	(m ²)
Longitudinal position of center of buoyancy	L_{CB}	(m)
Longitudinal position of center of buoyancy	L_{CB}	(m)
Bow shape coefficient		
Section shape coefficient		

For each displacement the equilibrium draft on even keel has been calculated, together with the necessary input variables [10] required by the Guldhammer Harvald method [29] to perform resistance prediction in calm waters. The propulsion coefficients have been corrected in magnitude as reported in [50].

Propeller thrust and torque were computed offline for different pitch settings by means of a viscous method and based on the knowledge of the geometrical features of the propeller. The calculated values were implemented in the matching code through the non dimensional thrust K_T and torque K_Q coefficients.

As reported in Fig. 1 a shaft generator is used for fulfilling on board auxiliary power demand. In order to optimise this feature the ship propulsion system has been set-up for working at fixed rpm using the pitch as control variable. Once the displacement, shaft rate of revolutions and vessel speed are fixed, the advance coefficient J is defined together with the non dimensional thrust coefficient according to the following equations:

$$J = \frac{V(1-w)}{nD}, \quad K_T = \frac{T}{\rho n^2 D^4} \quad (2)$$

where w is the wake factor, n is the propeller rate of revolution, D is the propeller diameter and T is the required thrust of the propeller. The engine-propeller matching code used in this work allows calculating the pitch ratio that provides the required thrust at the fixed shaft speed. Finally the delivered power P_d can be evaluated by means of the following quantities:

$$K_Q = \frac{Q}{\rho n^2 D^5}, \quad \eta_0 = \frac{J k_T}{2\pi K_Q} \quad (3)$$

A validation of the WBM model was performed based on the available measurements of delivered power at different displacement derived from model tests in calm

Table 5 White box model validation

$\Delta = 25000 \text{ t}$			
Speed	P_{dh} (KW)	P_{dn} (KW)	Error (%)
10	1452.0	1548.9	6.7
12	2589.0	2653.8	2.5
14	4478.0	4465.9	0.3
16	7761.9	7430.3	4.3
$\Delta = 30000 \text{ t}$			
10	1528.0	1637.6	7.2
12	2763.0	2805.3	1.5
14	4768.0	4818.1	1.1
16	18139.7	8139.1	0.0
$\Delta = 40000 \text{ t}$			
Speed	P_{dh} (KW)	P_{dn} (KW)	Error (%)
10	1768.0	1804.9	2.1
12	3100.0	3091.1	0.3
14	5196.0	5403.4	4.0
16	8760.2	9359.5	6.8
$\Delta = 50800 \text{ t}$			
Speed	P_{dh} (KW)	P_{dn} (KW)	Error (%)
10	1994.0	1815.5	9.0
12	3432.0	3106.3	9.5
14	5662.0	5458.5	3.6
16	9494.8	9546.7	0.5
$\Delta = 57100 \text{ t}$			
Speed	P_{dh} (KW)	P_{dn} (KW)	Error (%)
10	2162.0	1908.6	11.7
12	3690.0	3265.1	11.5
14	6089.0	5786.9	5.0
16	10323.4	10283.7	0.4

water. The measured (P_{dh}) and predicted (P_{dn}) delivered power, together with the absolute percentage error of the model, are reported in Table 5. The results obtained with the WBM model are in good agreement with measured values: thus, the model tool is able to derive a general representation of the relationship between vessel speed, displacement and delivered power in calm water scenarios.

For a generic couple of ship displacement Δ_i and speed V_i values, the WBM model evaluates the propeller rate of revolution n , which ensures the propulsion equilibrium between delivered and required thrust, and finally the associated fuel consumption. Starting from propeller torque, the engine brake power P_b is computed by the global

efficiency of the drivetrain and it is then possible to evaluate the corresponding specific fuel consumption.

5.2 Black Box Models

Machine Learning (ML) approaches play a central role in extracting information from raw data collected from ship data logging systems. The learning process for ML approaches usually consists of two phases: (i) during the training phase, a set of data is used to induce a model that best fits them, according to some criteria; (ii) the trained model is used for prediction and control of the real system (feed-forward phase).

As the authors are targeting a regression problem [74], the purpose is to find the best approximating function $h(\mathbf{x})$, where $h : \mathbb{R}^d \rightarrow \mathbb{R}$. During the training phase, the quality of the regressor $h(\mathbf{x})$ is measured according to a loss function $\ell(h(\mathbf{x}), y)$ [47], which calculates the discrepancy between the true and the estimated output (y and \hat{y}). The empirical error then computes the average discrepancy, reported by a model over \mathcal{D}_n :

$$\widehat{L}_n(h) = \frac{1}{n} \sum_{i=1}^n \ell(h(\mathbf{x}_i), y_i). \quad (4)$$

A simple criterium for selecting the final model during the training phase consists in choosing the approximating function that minimises the empirical error $\widehat{L}_n(h)$: this approach is known as Empirical Risk Minimisation (ERM) [74]. However, ERM is usually avoided² in ML as it leads to severely overfitting the model on the training dataset [74]. A more effective approach consists in the minimisation of a cost function where the tradeoff between accuracy on the training data and a measure of the complexity of the selected approximating function is implemented [72]:

$$h^* : \min_h \widehat{L}_n(h) + \lambda \mathcal{C}(h). \quad (5)$$

where $\mathcal{C}(\cdot)$ is a complexity measure which depends on the selected ML approach and λ is a hyperparameter that must be set a priori and regulates the trade-off between the overfitting tendency, related to the minimisation of the empirical error, and the underfitting tendency, related to the minimisation of $\mathcal{C}(\cdot)$. The optimal value for λ is problem-dependent, and tuning this hyperparameter is a non-trivial task [1] and will be faced later in this section.

The approaches exploited in this book chapter are: the Regularised Least Squares (RLS) [31], the Lasso Regression (LAR) [71], and the Random Forrest (RF) [5].

²Note that some techniques use ERM and then, in order to improve the performance of the method, a post processing approach is adopted (i.e. pruning for Decision Tree [59]).

In RLS, approximation functions are defined as

$$h(\mathbf{x}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}), \quad (6)$$

where a non-linear mapping $\boldsymbol{\phi} : \mathbb{R}^d \rightarrow \mathbb{R}^D$, $D \gg d$, is applied so that non-linearity is pursued while still coping with linear models.

For RLS, Problem (5) is configured as follows. The complexity of the approximation function is measured as

$$\mathcal{C}(h) = \|\mathbf{w}\|_2^2 \quad (7)$$

i.e. the Euclidean norm of the set of weights describing the regressor, which is a quite standard complexity measure in ML [72]. Regarding the loss function, the Mean Squared Error (MSE) loss is adopted:

$$\widehat{L}_n(h) = \frac{1}{n} \sum_{i=1}^n \ell(h(\mathbf{x}_i), y_i) = \frac{1}{n} \sum_{i=1}^n [h(\mathbf{x}_i) - y_i]^2. \quad (8)$$

Consequently, Problem (5) can be reformulated as:

$$\mathbf{w}^* : \min_{\mathbf{w}} \frac{1}{n} \sum_{i=1}^n [\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) - y_i]^2 + \lambda \|\mathbf{w}\|_2^2. \quad (9)$$

By exploiting the Representer Theorem [62], the solution h^* of the RLS Problem (9) can be expressed as a linear combination of the samples projected in the space defined by $\boldsymbol{\phi}$:

$$h^*(\mathbf{x}) = \sum_{i=1}^n \alpha_i \boldsymbol{\phi}(\mathbf{x}_i)^T \boldsymbol{\phi}(\mathbf{x}). \quad (10)$$

It is worth underlining that, according to the kernel trick [61], it is possible to reformulate $h^*(\mathbf{x})$ without an explicit knowledge of $\boldsymbol{\phi}$ by using a proper kernel function $K(\mathbf{x}_i, \mathbf{x}) = \boldsymbol{\phi}(\mathbf{x}_i)^T \boldsymbol{\phi}(\mathbf{x})$:

$$h^*(\mathbf{x}) = \sum_{i=1}^n \alpha_i K(\mathbf{x}_i, \mathbf{x}). \quad (11)$$

Of the several kernel functions which can be found in literature [15], the Gaussian kernel is often used as it enables learning every possible function [56]:

$$K(\mathbf{x}_i, \mathbf{x}_j) = e^{-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|_2^2}, \quad (12)$$

where γ is an hyperparameter which regulates the non-linearity of the solution [56] and must be set a priori, analogously to λ . Small values of γ lead the optimisation to converge to simpler functions $h(x)$ (note that for $\gamma \rightarrow 0$ the optimisation converges to a linear regressor), while high values of γ allow higher complexity of $h(x)$.

Finally, the RLS Problem (9) can be reformulated by exploiting kernels:

$$\alpha^* : \min_{\alpha} \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^n \alpha_j K(\mathbf{x}_j, \mathbf{x}_i) - y_i \right]^2 + \lambda \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j K(\mathbf{x}_j, \mathbf{x}_i). \quad (13)$$

Given $\mathbf{y} = [y_1, \dots, y_n]^T$, $\alpha = [\alpha_1, \dots, \alpha_n]^T$, the matrix K such that $K_{i,j} = K_{j,i} = K(\mathbf{x}_j, \mathbf{x}_i)$, and the Identity matrix $I \in \mathbb{R}^{n \times n}$, a matrix-based formulation of Problem (13) can be obtained:

$$\alpha^* : \min_{\alpha} \frac{1}{n} \|K\alpha - \mathbf{y}\|_2^2 + \lambda \alpha^T K \alpha \quad (14)$$

By setting the derivative with respect to α equal to zero, α can be found by solving the following linear system:

$$(K + n\lambda I) \alpha^* = \mathbf{y}. \quad (15)$$

Effective solvers have been developed throughout the years, allowing to efficiently solve the problem of Eq. (15) even when very large sets of training data are available [80].

In LAR, instead, approximation functions are defined as

$$h(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b, \quad (16)$$

which are linear functions in the original space \mathbb{R}^d .

For LAR, Problem (5) is configured as follows. The complexity of the approximation function is measured as

$$\mathcal{C}(h) = \|\mathbf{w}\|_1 \quad (17)$$

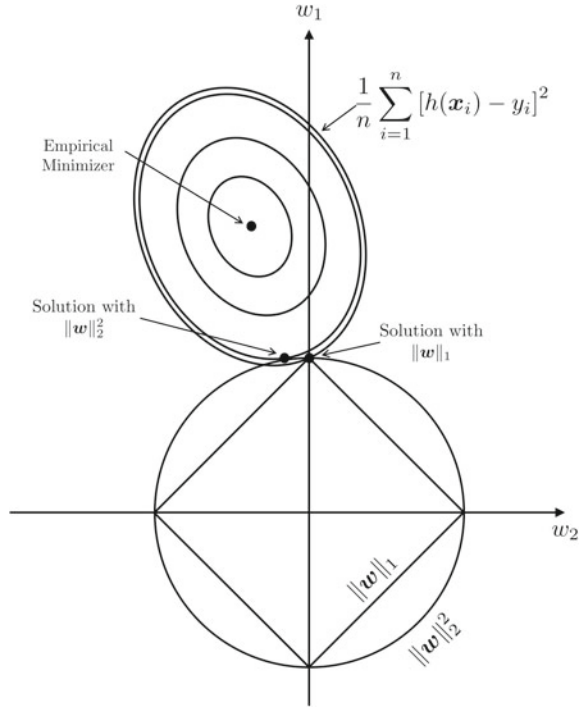
i.e. the Manhattan norm of the set of weights describing the regressor [71].

Regarding the loss function, the Mean Squared Error (MSE) loss is again adopted. Consequently, Problem (5) can be reformulated as:

$$\mathbf{w}^* : \min_{\mathbf{w}} \frac{1}{n} \sum_{i=1}^n [\mathbf{w}^T \phi(\mathbf{x}) - y_i]^2 + \lambda \|\mathbf{w}\|_1. \quad (18)$$

As depicted in Fig. 7 the Manhattan norm is quite different from the Euclidean one since it allows increasing the sparsity of the solution. In other words the solution will

Fig. 7 Manhattan norm Versus euclidean norm



tend to fall on the edge of the square, forcing some weights of \mathbf{w} to be zero. Hence, the Manhattan norm allows both regularising the function and discarding features that are not sufficiently relevant to the model. This property is particularly useful in the feature selection process [71].

Two main approaches can be used to compute the solutions of Problem (18): the LARS algorithms [19] and the pathwise coordinate descent [23]. In this book chapter, the LARS algorithm is exploited because of its straight-forward implementation [19].

The performance of RLS (or LAR) models depends on the quality of the hyperparameters tuning procedure. As highlighted while presenting this approach, the parameters α^* , $\hat{\alpha}^*$, and $\check{\alpha}^*$ (or \mathbf{w}) result from an optimisation procedure which requires the a priori setting of the tuples of hyperparameters (λ, γ) (or λ). The phase in which the problem of selecting the best value of the hyperparameter is addressed is called model selection phase [1]. The most effective model selection approaches consist in performing an exhaustive hyperparameters grid search: the optimisation problem for RLS (or LAR) is solved several times for different values of γ and λ , and the best pair of hyperparameters is chosen according to some criteria.

For the optimal choice of the hyperparameters γ and λ , in this book chapter the authors exploit the Bootstrap technique (BOO) [1]. This technique represents an improvement of the well-known k -Fold Cross Validation (KCV) [44] where the original dataset is split into k independent subsets (namely, the folds), each one

consisting of n/k samples: $(k - 1)$ parts are used, in turn, as a training set, and the remaining fold is exploited as a validation set. The procedure is iterated k times.

The standard Bootstrap [1] method is a pure resampling technique: at each j -th step, a training set $\mathcal{D}_{\text{TR}}^j$, with the same cardinality of the original one, is built by sampling the patterns in \mathcal{D}_n with replacement. The remaining data $\mathcal{D}_{\text{VL}}^j$, which consists, on average, of approximately 36.8% of the original dataset, are used as validation set. The procedure is then repeated several times $N_B \in [1, \binom{2n-1}{n}]$ in order to obtain statistically sound results [1].

According to the Bootstrap technique, at each j -th step the available dataset \mathcal{D}_n is split in two sets:

- A training Set: $\mathcal{D}_{\text{TR}}^j$
- A validation Set: $\mathcal{D}_{\text{VL}}^j$

In order to select the best pair of hyperparameters (λ^*, γ^*) (or λ^*) among all the available ones $\mathcal{G} = \{(\lambda_1, \gamma_1), (\lambda_2, \gamma_2), \dots\}$ (or $\mathcal{G} = \{\lambda_1, \lambda_2, \dots\}$) for the algorithm for RLS (or LAR) the following optimisation procedure is required:

- for each $\mathcal{D}_{\text{TR}}^j$ and for each tuple (λ_i, γ_i) (or λ_i) with $i \in \{1, 2, \dots\}$ the optimisation problem of Eq. 15 (or Eq. 18) is solved and the solution $h_j^i(\mathbf{x})$ is found
- using the validation set $\mathcal{D}_{\text{VL}}^j$ for searching the (λ^*, γ^*) (or $\lambda^*) \in \mathcal{G}$

$$\begin{aligned} (\lambda^*, \gamma^*) \\ \text{or } \lambda^* = \arg \min_{\substack{\{(\lambda_1, \gamma_1), \dots, (\lambda_i, \gamma_i), \dots\} \\ \text{or } \{\lambda_1, \dots, \lambda_i, \dots\}}} \frac{1}{N_B} \sum_{j=1}^{N_B} \frac{1}{|\mathcal{D}_{\text{VL}}^j|} \sum_{(\mathbf{x}, y) \in \mathcal{D}_{\text{VL}}^j} [h(\mathbf{x}) - y]^2. \end{aligned} \quad (19)$$

Once the best tuple is found, the final model is trained on the whole set \mathcal{D}_n by running the learning procedure with the best values of the hyperparameters [1].

Another learning algorithm tested for building the BBM is the Random Forest (RF) [5]. Random Forests grows many regression trees. To classify a new object from an input vector each of the trees of the forest is applied to the vector. Each tree gives an output and the forest chooses the mode of the votes (over all the trees in the forest). Each single tree is grown by following this procedure: (I) n samples are sampled (with replacement) from the original \mathcal{D}_n , (II) $d' \ll d$ features are chosen randomly out of the d and the best split on these d' is used to split the node, (III) each tree is grown to the largest possible extent, without any pruning. In the original paper [5] it was shown that the forest error rate depends on two elements: the correlation between any couples of trees in the forest (increasing the correlation increases the forest error rate) and the strength of each individual tree in the forest (reducing the error rate of each tree decreases the forest error rate). Reducing d' reduces both the correlation and the strength. Increasing it increases both. Somewhere in between is an optimal range of d' - usually quite wide so this is not usually considered as an hyperparameter. Note that, since we used a bootstrap procedure by sampling n

samples with replacement from the original \mathcal{D}_n , we can use the error on the remaining part of the data (this is called out-of-bag error) to chose the best d' .

5.3 Gray Box Models

GBMs are a combination of a WBM and BBMs. This requires to modify the BBMs as defined in the previous section in a way to include the mechanistic knowledge of the system. Two approaches are tested and compared in this book chapter:

- a Naive approach (N-GBM) where the output of the WBM is used as a new feature that the BBM can use for training the model.
- an Advanced approach (A-GBM) where the regularisation process is changed in order to include some a-priori information [1].

In the N-GBM case, the WBM can be seen as a function of the input \mathbf{x} . The WBM, that we call here $h_{\text{WBM}}(\mathbf{x})$, allows the creation of a new dataset:

$$\mathcal{D}_n^{\text{WBM}, \mathcal{X}} = \left\{ \left(\begin{bmatrix} \mathbf{x}_1 \\ h_{\text{WBM}}(\mathbf{x}_1) \end{bmatrix}, y_1 \right), \dots, \left(\begin{bmatrix} \mathbf{x}_n \\ h_{\text{WBM}}(\mathbf{x}_n) \end{bmatrix}, y_n \right) \right\}$$

Based on this new dataset a BBM can be generated $h_{\text{BBM}} \left(\left[\mathbf{x}^T | h_{\text{WBM}}(\mathbf{x}) \right]^T \right)$.

According to this approach, every run of the GBM requires an initial run of the WBM in order to compute its output $h_{\text{WBM}}(\mathbf{x})$, which allows evaluating the model $h_{\text{BBM}} \left(\left[\mathbf{x}^T | h_{\text{WBM}}(\mathbf{x}) \right]^T \right)$. This is the simplest approach for including new information into the learning process. Note that with this approach any of the previously cited BBMs (e.g. RLS, LAR or RM) can be used for building the corresponding N-GBM.

In the A-GBM case the WBM part of the model is assumed to be included in the \mathbf{w} vector:

$$h_{\text{WBM}}(\mathbf{x}) = \mathbf{w}_{\text{WBM}}^T \boldsymbol{\phi}(\mathbf{x}), \quad (20)$$

According to [1], the regularisation process of Eq. (9) is modified to:

$$\mathbf{w}^*: \min_{\mathbf{w}} \frac{1}{n} \sum_{i=1}^n [\mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) - y_i]^2 + \lambda \|\mathbf{w} - \mathbf{w}_{\text{WBM}}\|_2^2. \quad (21)$$

It is possible to prove that by exploiting the kernel trick the solution to this problem can be rewritten as:

$$h^*(\mathbf{x}) = h_{\text{WBM}}(\mathbf{x}) + \sum_{i=1}^n \alpha_i^* K(\mathbf{x}_i, \mathbf{x}). \quad (22)$$

where

$$\boldsymbol{\alpha}^* : \min_{\boldsymbol{\alpha}} \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^n \alpha_j K(\mathbf{x}_j, \mathbf{x}_i) + h_{\text{WBM}}(\mathbf{x}_i) - y_i \right]^2 + \lambda \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j K(\mathbf{x}_j, \mathbf{x}_i), \quad (23)$$

The solution to this problem can be computed by solving the following linear system:

$$(K + n\lambda I) \boldsymbol{\alpha}^* = \mathbf{y} - \mathbf{h}_{\text{WBM}}, \quad (24)$$

where $\mathbf{h}_{\text{WBM}} = [h_{\text{WBM}}(\mathbf{x}_1), \dots, h_{\text{WBM}}(\mathbf{x}_n)]^T$. Note that the solution does not depend on the form of $h_{\text{WBM}}(\mathbf{x})$ so that any WBM can be used as $h_{\text{WBM}}(\mathbf{x})$.

Another possible way of achieving the same solution of the problem of Eq. (24) is to create a new dataset:

$$\mathcal{D}_n^{\text{WBM}, \mathcal{E}} = \{(\mathbf{x}_1, y_1 - h_{\text{WBM}}(\mathbf{x}_1)), \dots, (\mathbf{x}_n, y_n - h_{\text{WBM}}(\mathbf{x}_n))\}$$

where the target is no longer the true label y but the true label minus the hint given by the a priori information included in $h_{\text{WBM}}(\mathbf{x})$. This means finding a BBM that minimises the error of the WBM prediction.

It should be noted that the A-GBM is more theoretically justified in the regularisation context while the N-GBM is more intuitive since all the available knowledge is given as input to the BBM learning process. From a probabilistic point of view, the A-GMB changes the $\mathbb{P}(y|\mathbf{x})$ while the N-GBM modifies the whole joint probability $\mathbb{P}(y, \mathbf{x})$, hence deeply influencing the nature of the problem.

The (λ, γ) for RLS, the λ for LAR and d' for RF of the N-GBM and A-GBM are tuned with the BOO as described for the BBM, since both N-GBM and A-GBM basically require to build a BBM over a modified training set.

5.4 Model Validation

The WBM was validated using the data described in Sect. 3 versus propeller shaft power, shaft torque, and total fuel consumption. The results of the validation are presented in Table 6. The results show that the WBM does not show sufficient accuracy when compared with operational measurements. The inability of the model to take into account the influence of the sea state (i.e. wind and waves) on the required propulsion power is considered to be the largest source of error for this model.

The BBMs built according to the RLS, LAR and RF methods were validated versus the same dataset as for the WBM validation procedure. However, in the case of the BBMs the \mathcal{D}_n was divided in two sets \mathcal{L}_{n_l} and \mathcal{T}_{n_t} respectively for learning and test. The two sets were defined so that $\mathcal{D}_n = \mathcal{L}_{n_l} \cup \mathcal{T}_{n_t}$ and $\mathcal{L}_{n_l} \cap \mathcal{T}_{n_t} = \emptyset$ in order to maintain the independence of the two sets.

Table 6 Indexes of performance of the WBM in predicting the shaft power, shaft torque, and fuel consumption

Shaft power					
MAE (KW)	MAPE (%)	MSE (KW ²)	NMSE	REP (%)	PPMCC
7.69e+02	17.85	1.00e+06	1.13	23.59	0.65
Shaft torque					
MAE (Nm)	MAPE (%)	MSE (N ² m ²)	NMSE	REP (%)	PPMCC
6.54e+01	18.13	6.92e+03	0.94	22.01	0.22
Fuel consumption					
MAE ($\frac{g}{kWh}$)	MAPE (%)	MSE ($\frac{g^2}{kWh^2}$)	NMSE	REP (%)	PPMCC
5.14e-02	20.95	3.94e-03	1.98	25.40	0.63

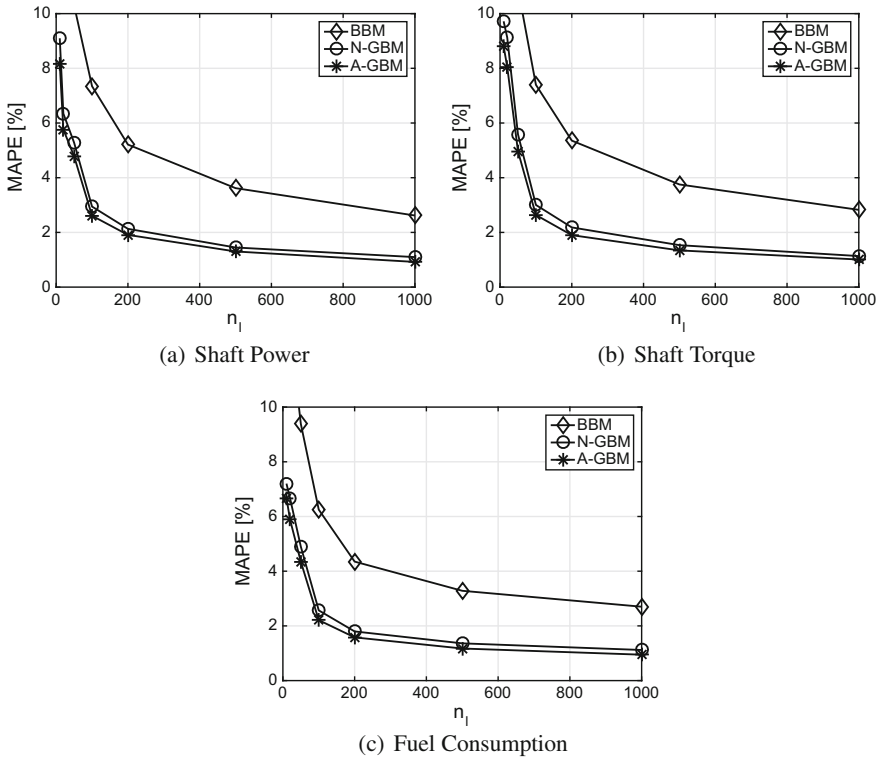


Fig. 8 Shaft Power, Shaft Torque, and Fuel Consumption MAPE of the BBM, N-GBM and A-GBM for RLS and different n_l

The process of splitting the full dataset in a learning set and test set is repeated 30 times in order to obtain statistical relevant results. We always underline in bold the best results which are statistically significant [25]. The results are reported for different sizes of \mathcal{L}_{n_l} with $n_l \in \{10, 20, 50, 100, 200, 500, 1000, 2000, 5000\}$. The optimisation procedure is repeated for different values of both hyperparameters (γ and λ), where their values are taken based on a 60 points equally spaced in logarithmic scale in the range $[10^{-6}, 10^3]$ and the best set of hyperparameters is selected according to the BOO (Sect. 5.2). The same has been done for d' in RF.

Also in the case of the GBM, analogously to the procedure adopted for the BBM, the original dataset \mathcal{D}_n is divided in two sets \mathcal{L}_{n_l} and \mathcal{T}_{n_l} and λ , γ and d' are chosen according to the BOO procedure.

The entire set of results is not reported here because of space constraints but it can be retrieved in the technical report available at <http://www.smartlab.ws/TR.pdf>

From the results it is possible to note that the WBM, as expected, has the lowest performance in terms of prediction accuracy. On the other hand, the GBMs outperform the BBMs by a smaller percentage. The MAPE of the BBM, N-GBM and A-GBM for different values of n_l are reported in Figs. 8, 9, and 10.

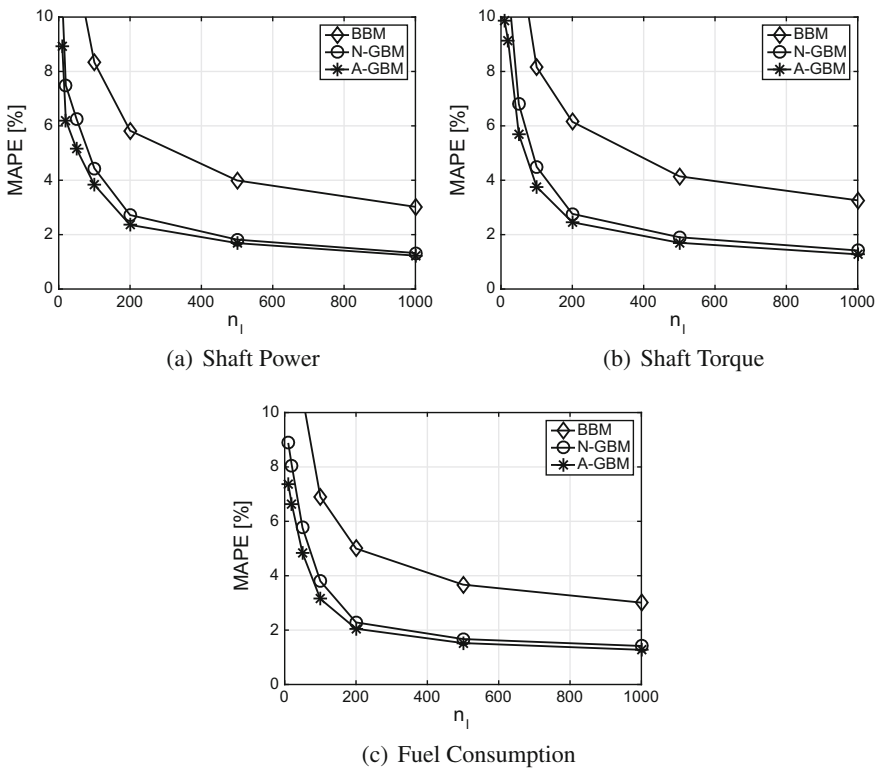


Fig. 9 Shaft Power, Shaft Torque, and Fuel Consumption MAPE of the BBM, N-GBM and A-GBM for LAR and different n_l

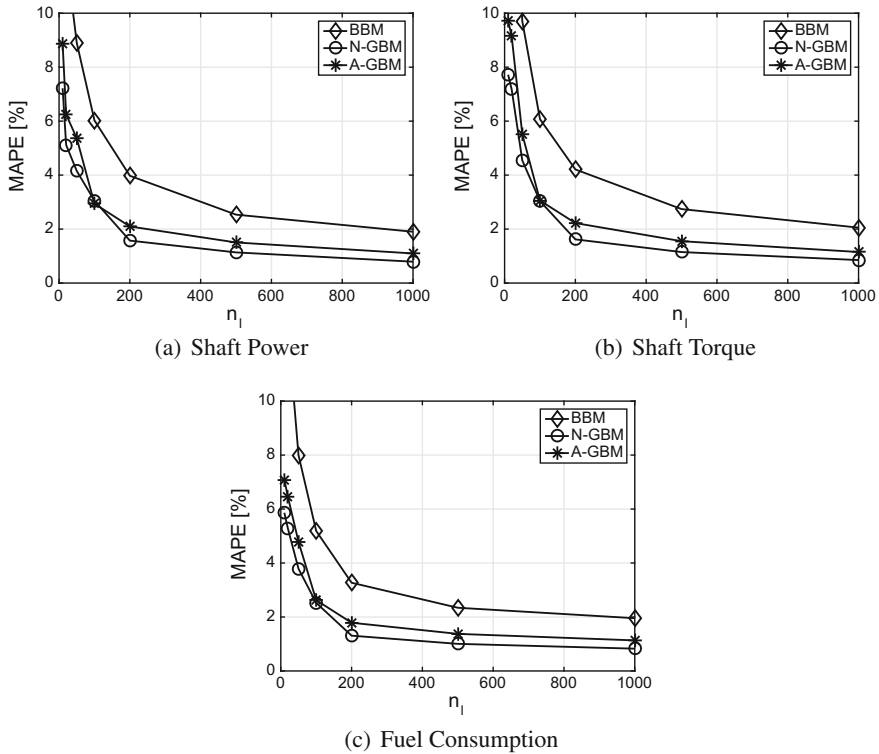


Fig. 10 Shaft Power, Shaft Torque, and Fuel Consumption MAPE of the BBM, N-GBM and A-GBM for RF and different n_l

From the results of Figs. 8, 9, and 10 it is possible to note how the WBM, even if not so accurate, can help the GBM in obtaining higher accuracy, with respect to the BBM, by using almost half of the data given a required accuracy. This is a critical issue in real word applications where the collection of labeled data can be expensive or at least requires a long period of in-service operational time of the vessel [11].

6 Feature Selection

Once a model is built and has been confirmed to be a sufficiently accurate representation of the real system of interest, it can be interesting to investigate how the model \mathfrak{M} is affected by the different features that have been used in the model identification phase.

In data analytics this procedure is called feature selection or feature ranking [7, 22, 30, 36, 79]. This process allows detecting if the importance of those features, that are known to be relevant from a theoretical perspective, is appropriately described by \mathfrak{M} .

The failure of the statistical model to properly account for the relevant features might indicate poor quality in the measurements. Feature selection therefore represents an important step of model verification, since the proposed model \mathfrak{M} should generate results consistently with the available knowledge of the physical system under exam. This is particularly important in the case of BBM (and, to a more limited extent, for GBM), since they do not make use of any mechanistic knowledge of the system and might therefore lead to non-physical results (e.g. mass or energy unbalances). Feature selection also allows checking the statistical robustness of the employed methods.

In this book chapter, three different methods for feature ranking are applied:

- Brute Force Method (BFM), which searches for the optimal solution. This is the most accurate method but also the most computationally expensive (see Sect. 6.1) [22, 27].
- Regularisation Based Method (RBM) which works by building the BBM which automatically discarding the features that do not significantly contribute to the model output (for example by building an ad-hoc regularisers [16, 22, 53, 55, 66, 81, 82]). In this book chapter, the Lasso Regularisation technique was used (see Sect. 6.2).
- Random Forest based method (RFM) uses a combination of Decision Tree methods together with the permutation test [28] in order to perform the selection and the ranking of the features [22, 43, 70].

6.1 Brute Force Method

According to the Brute Force method (BFM) for feature selection, the k most important features of the model can be identified as follows:

- a first version of the model \mathfrak{M} including all the available features is built. The full model is tested against a test set $\widehat{L}_{\text{Test}}$;
- for a given k , a set of new models is built for all possible configurations including feature k . For every possible configuration, which are $\binom{d}{k}$, a new model \mathfrak{M}^j is built where $j \in \left\{1, \dots, \binom{d}{k}\right\}$ together with its error on the test set $\widehat{L}_{\text{Test}}^j$;
- the smaller is the difference between $\widehat{L}_{\text{Test}}^j$ and $\widehat{L}_{\text{Test}}$, the greater is the importance of that set of features.

Given its high computational demands, this approach is not feasible for $d > 15 \div 20$. A solution for reducing the required computational time is to adopt a greedy procedure:

- a first version of the model \mathfrak{M} including all the available features is built. The full model is tested against a test set $\widehat{L}_{\text{Test}}$;
- given a feature j_1 , the model \mathfrak{M} is built which only includes that feature. The error against the test set ($\widehat{L}_{\text{Test}}^{j_1}$) can now be calculated;
- the same procedure is performed for each feature $j_1 \in \{1, \dots, d\}$;

- the smaller is the difference between $\widehat{L}_{\text{Test}}^{j_1}$ and $\widehat{L}_{\text{Test}}$ the grater is the importance of the features j_1

$$j_1^* = \arg \min_{j_1 \in \{1, \dots, d\}} \widehat{L}_{\text{Test}} - \widehat{L}_{\text{Test}}^{j_1} \quad (25)$$

- this procedure is repeated by adding to j_1^* all the other features one at the time for finding the second most important feature $j_2^* \in \{1, \dots, d\} \setminus j_1^*$. This operation is repeated until the required size (k) of the ranking is achieved.

Greedy methods are more time efficient compared to brute force methods, but do not ensure the full correctness of the result.

In this book chapter, several different models were proposed and are here tested for feature ranking. These models are: BBM, N-GBM and A-GBM with RLS, LAR and RF for a total of nine possibilities. It should be noted that for the N-GBM there is another feature which is the WBM (see Table 3).

6.2 Regularisation Based Method

The brute force method is a quite powerful approach but it requires a significant computational effort. The Lasso Regression can be used for ranking the importance of the features with lower computational demand.

However, the results of the Lasso Regression method are strongly influenced by the training dataset and by the choice of the hyperparameters used in the learning phase [53, 81, 82]. For this reason given the best value λ^* of the hyperparameter selected with the BOO procedure another bootstrap procedure is applied in order to improve the reliability of the feature selection method: n samples are extracted from \mathcal{D}_n , the model is built with LAR and λ^* and the features are selected. The bootstrap is repeated several times and features are ranked based on how many times each feature is selected as important by the LAR method [53].

In this work, the LAR method for feature selection was used in three different kind of models (BBM, N-GBM and A-GBM). Similarly to the case of Brute Force Methods, in the N-GBM case the WBM represents an additional feature (see Table 3).

6.3 Random Forest Based Method

In addition to its use for regression models, the Random Forest (RF) method can also be used to perform a very stable Feature Selection procedure. The procedure can be described as follows: in every tree grown in the forest the error on the out-of-bag must be kept. Then a random permutation of the values of variable j must be performed in the samples of the out-of-bag and the error on the out-of-bag must be kept again. Subtract the error on the untouched out-of-bag data with the error over the permuted

out-of-bag samples. The average of this value over all the trees in the forest is the raw importance score for variable j . This approach is inspired by the permutation test [28] which is quite used in literature, is computationally inexpensive in the case of Random Forest, and has shown to be quite effective in real world applications [17, 77]. The results for the RF feature selection method are reported for all the models (BBM, N-BBM and A-GBM).

6.4 Results

In Table 7 all the results of the feature selection method are reported. Authors decided to provide just the seven most informative features not to compromise the readability of the tables. From the tables it is possible to draw the following considerations:

- all methods identify the same variables as the most relevant for the model, thus confirming the validity of the modelling procedure. This also allows to trust the reliability of the information contained historical data.
- the BF methods are the most stable, closely followed by the RF methods.
- the WBM is always among the seven most important features for GBMs. This suggests that the N-GBM is able to take into account the information generated by the WBM and use it appropriately, confirming the results of the previous section which underlined the improved performance of GBMs compared to BBMs.

From a physical point of view the results of the feature selection identify the propeller pitch (both setpoint and feedback) and the ship speed (both GPS and LOG) as the most important variables for the prediction, which is what to be expected from this type of ship propulsion system. Propeller speed is not among the most important features, as it is normally kept constant during ship operations and therefore has very limited impact from a modelling perspective. The ship draft (fore and aft) are normally selected as important variables (5th-6th), which also reflects physical expectations from the system as the draft influences both ship resistance and, to a minor extent, propeller performance. As expected the shaft generator power, for this propulsion plants configuration, plays an important role for the prediction. In addition to this, some variables that could be expected to contribute significantly to the overall performance are missing. In particular, wind speed and direction are generally used for estimating the impact of the sea state, but are not included among the five most relevant features by any feature selection method. This suggests that either the sea state has a less significant impact on the ship's fuel consumption compared to what originally expected, or that wind speed and direction are not appropriate predictors for modelling this type of effects, contrarily to what often assumed in relevant literature. One possible additional explanation to the absence of wind speed and direction from the important variables is that the influence of the sea state is already accounted for by the propeller pitch ratio, which is expected to vary as a consequence of both ship speed and ship added resistance. As matter of fact in order to keep constant speed

Table 7 Feature ranking with the BF, LAR and RF with just the seven most informative features

BFM		RBM							RFM							
RLS		LAR				RF										
Ranking	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	
Shaft rpm																
1°	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	
2°	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	
3°	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
4°	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	
5°	21	WBM	21	21	WBM	21	21	WBM	21	9	WBM	21	21	WBM	21	
6°	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	
7°	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	
Shaft Power																
1°	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	
2°	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
3°	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	
4°	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	
5°	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	
6°	19	WBM	19	19	WBM	19	19	WBM	19	16	10	19	19	WBM	19	
7°	10	10	10	10	10	10	10	10	10	10	WBM	9	10	10	10	
Fuel Consumption																
1°	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	
2°	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	
3°	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	

(continued)

Table 7 (continued)

		BFM				RBM				RFM					
		RLS		LAR		RF									
Ranking	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM	BBM	N-GBM	A-GBM
4°	21	21	21	21	21	21	21	21	21	21	9	16	21	21	21
5°	9	9	9	9	9	9	9	9	9	9	16	9	9	9	9
6°	16	WBM	16	16	WBM	16	16	WBM	16	16	WBM	21	16	WBM	16
7°	10	10	10	10	10	10	10	10	10	27	27	27	10	10	10

profile, the on board automation system should be designed to change the pitch settings and the fuel consumption rate to take into account time domain variation of boundary conditions such as wind and sea state conditions. Under this assumption the relevant information about added resistance and wind intensity are already included in the propeller pitch ratio.

7 Using Machine Learning for Operational Energy Savings: Trim Optimisation

Of all the models proposed in the previous part of this book chapter, the N-GBM based on RF features the best accuracy properties and best physical plausibility and is therefore used for the trim optimisation problem. In order to meet the requirements expressed in Sect. 4.3 the following is considered:

- Variables that are influenced by the trim, such as propeller power and torque, were excluded from the model (see Table 3).
- For each pair of ship speed and displacement, the trim is only allowed to vary in the range observed from the available dataset, extended by $\delta\%$. This allows, for every pair, to limit the extrapolation and therefore to ensure additional reliability of the optimisation results.

In Table 8 the Fuel Consumption percentage reduction with the trim Optimisation technique is reported for different values of δ . As expected, the optimisation procedure always leads to a reduction in fuel consumption. The improvement that can be achieved via trim optimisation increases when δ is increased, although this tendency seems to stabilise for $\delta > 5\%$.

According to the results of this model, improvements exceeding 2% in fuel consumption can be achieved by applying the model for trim optimisation to the selected vessel. It should be noted that trim optimisation can be performed at near to zero cost on board, since it does not require the installation of any additional equipment. Future work in this area will include testing trim optimisation system here proposed on a real vessel, in order to check the validity of the model and the performance of the optimisation tool.

Table 8 Fuel consumption percentage reduction with the trim optimisation technique

$\delta(\%)$	% reduction
0	0.52 ± 0.12
1	1.45 ± 0.32
2	1.72 ± 0.51
5	2.22 ± 0.67
10	2.30 ± 0.64

8 Summary

This chapter focused on the utilisation of methods of Machine Learning for making ship operations more sustainable. Shipping is today facing large challenges in terms of its impact on the climate, and the reduction of CO₂ emissions that are expected to be achieved in the future will require a significant effort.

The achievement of such goals will require, among others, to improve today's ability to accurately model and predict the influence of environmental and operational variables on ship performance, and in particular on the fuel consumption of the ship. In this chapter, alongside with the white-box models commonly used today in this industry, black and gray box models were introduced as modelling approaches that can improve the accuracy of the prediction by making use of extensive measured data from ship operations. The regularised least squares, Lasso Regression and Random forest methods for the construction of black box models were proposed. In addition, two different types of gray, "hybrid" modelling approaches combining elements of white and black box models were also presented: the naive, and the advanced approach. Finally, feature selection methods were introduced, that can be used for testing the physical consistency of black and gray box models.

The book chapter was concluded with the application of the proposed methods to a case study, a chemical tanker, with the aim of testing their ability of predicting fuel consumption and of optimising the trim of the vessel. The results of this application case confirmed the superiority of statistical methods over mechanistic models in their ability of accurately predict the performance of the vessel, and highlighted that gray box models, although improving the performance of black box models only marginally, show an increased predictive ability with small sizes of the training dataset. The application of a naive-gray box model to the problem of trim optimisation allowed identifying the possibility of decreasing fuel consumption by up to 2.3% without the need of installing further equipment on board.

References

1. Anguita, D., Ghio, A., Oneto, L., Ridella, S.: In-sample and out-of-sample model selection and error estimation for support vector machines. *IEEE Trans. Neural Netw. Learn. Syst.* **23**(9), 1390–1406 (2012)
2. Armstrong, V.N.: Vessel optimisation for low carbon shipping. *Ocean Eng.* **73**, 195–207 (2013)
3. Baldi, F., Johnson, H., Gabriellii, C., Andersson, K.: Energy and exergy analysis of ship energy systems-the case study of a chemical tanker. In: 27th ECOS, International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (2014)
4. Basin, D.W.T.M., Todd, F.H.: Series 60 Methodical Experiments with Models of Single-Screw Merchant Ships. Washington (1964)
5. Breiman, L.: Random forests. *Mach. Learn.* **45**(1), 5–32 (2001)
6. Buhaug, O., Corbett, J.J., Endersen, O., Eyring, V., Faber, J., Hanayama, S., Lee, D.S., Lee, D., Lindstad, H., Markowska, A.Z., Mjelde, A., Nilsen, J., Palsson, C., Winebrake, J.J., Wu, W., Yoshida, K.: Second IMO GHG study 2009. Technical reports, International Maritime Organization (IMO) (2009)

7. Chang, Y.W., Lin, C.J.: Feature ranking using linear svm. *Causation Predict. Chall. Chall. Mach. Learn.* **2**, 47 (2008)
8. Coraddu, A., Figari, M., Savio, S., Villa, D., Orlandi, A.: Integration of seakeeping and powering computational techniques with meteo-marine forecasting data for in-service ship energy assessment. In: *Developments in Maritime Transportation and Exploitation of Sea Resources* (2013)
9. Coraddu, A., Gaggero, S., Figari, M., Villa, D.: A new approach in engine-propeller matching. In: *Sustainable Maritime Transportation and Exploitation of Sea Resources*, vol. 1, pp. 631–637. CRC Press —Taylor and Francis Group (2011)
10. Coraddu, A., Gualeni, P., Villa, D.: Investigation about wave profile effects on ship sability. *IMAM 2011 international maritime association of the mediterranean - sustainable maritime transportation and exploration of the sea. Resources* **1**, 143–149 (2011)
11. Coraddu, A., Oneto, L., Baldi, F., Anguita, D.: A ship efficiency forecast based on sensors data collection: improving numerical models through data analytics. In: *OCEANS* (2015)
12. Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., Figari, M.: Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers, Part M: J. Eng. Marit. Environ.* doi:[10.1177/1475090214540874](https://doi.org/10.1177/1475090214540874) (2014)
13. Corbett, J.J., Koehler, H.W.: Updated emissions from ocean shipping. *J. Geophys. Res. Atmos.* **108**(D20) (2003)
14. Cox, D.R.: *Principles of Statistical Inference*. Cambridge University Press (2006)
15. Cristianini, N., Shawe-Taylor, J.: *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge University Press (2000)
16. De Mol, C., De Vito, E., Rosasco, L.: Elastic-net regularization in learning theory. *J. Complex.* **25**(2), 201–230 (2009)
17. Deng, H., Runger, G., Tuv, E.: Bias of importance measures for multi-valued attributes and solutions. *Artif. Neural Netw. Mach. Learn. ICANN* **2011**, 293–300 (2011)
18. Devanney, J.: The impact of the energy efficiency design index on very large crude carrier design and CO₂ emissions. *Ships Offshore Struct.* **6**(4), 355–368 (2011)
19. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R.: Least angle regression. *Ann. Stat.* **32**(2), 407–499 (2004)
20. European Commission: *Integrating maritime transport emissions in the EU's greenhouse gas reduction policies* (2013)
21. Evans, J.R., Lindner, C.H.: Business analytics: the next frontier for decision sciences. *Decis. Line* **43**(2), 4–6 (2012)
22. Friedman, J., Hastie, T., Tibshirani, R.: *The Elements of Statistical Learning*. Springer series in statistics Springer, Berlin (2001)
23. Friedman, J., Hastie, T., Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**(1), 1 (2010)
24. Fumeo, E., Oneto, L., Anguita, D.: Condition based maintenance in railway transportation systems based on big data streaming analysis. In: *INNS Conference on Big Data* (2015)
25. García, S., Fernández, A., Luengo, J., Herrera, F.: Advanced nonparametric tests for multiple comparisons in the design of experiments in computational intelligence and data mining: experimental analysis of power. *Inf. Sci.* **180**(10), 2044–2064 (2010)
26. Ghelardoni, L., Ghio, A., Anguita, D.: Energy load forecasting using empirical mode decomposition and support vector regression. *IEEE Trans. Smart Grid* **4**(1), 549–556 (2013)
27. Gieseke, F., Polsterer, K.L., Oancea, C.E., Igel, C.: Speedy greedy feature selection: better redshift estimation via massive parallelism. In: *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning* (2014)
28. Good, P.: *Permutation Tests: A Practical Guide To Resampling Methods For Testing Hypotheses*. Springer Science & Business Media (2013)
29. Gulddhammer, H., Harvald, S.A.: *Ship Resistance: Effect of Form and Principal Dimensions*. Akademisk Forlag (1974)

30. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *J. Mach. Learn. Res.* **3**, 1157–1182 (2003)
31. Györfi, L.: *A Distribution-free Theory of Nonparametric Regression*. Springer (2002)
32. Haas, P.J., Maglio, P.P., Selinger, P.G., Tan, W.C.: Data is dead... without what-if models. In: *International Conference on Very Large Database* (2011)
33. Hochkirch, K., Mallol, B.: On the importance of fullscale cfd simulations for ships. In: *International Conference on Computer Applications and Information Technology in the Maritime Industries* (2013)
34. Holtrop, J.: A statistical re-analysis of resistance and propulsion data. *Int. Shipbuild. Prog.* **31**(363), 272–276 (1984)
35. Holtrop, J., Mennen, G.G.: An approximate power prediction method. *Int. Shipbuild. Prog.* **29**, 166–171 (1982)
36. Hong, S.J.: Use of contextual information for feature ranking and discretization. *IEEE Trans. Knowl. Data Eng.* **9**(5), 718–730 (1997)
37. Howison, S.: *Practical Applied Mathematics: Modelling, Analysis, Approximation*. 38. Cambridge University Press (2005)
38. Iakovatos, M.N., Liarokapis, D.E., Tzabiras, G.D.: Experimental investigation of the trim influence on the resistance characteristics of five ship models. In: *Developments in Maritime Transportation and Exploitation of Sea Resources—Proceedings of IMAM 2013, 15th International Congress of the International Maritime Association of the Mediterranean* (2014)
39. Jafarzadeh, S., Utne, I.B.: A framework to bridge the energy efficiency gap in shipping. *Energy* **69**, 603–612 (2014)
40. Jardine, A.K., Lin, D., Banjevic, D.: A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech. Syst. Signal Process.* **20**(7), 1483–1510 (2006)
41. Josephson, J.R., Josephson, S.G.: *Abductive Inference: Computation, Philosophy, Technology*. Cambridge University Press (1996)
42. Khor, Y.S., Xiao, Q.: CFD simulations of the effects of fouling and antifouling. *Ocean Eng.* **38**(10), 1065–1079 (2011)
43. Kohavi, R., John, G.H.: Wrappers for feature subset selection. *Artif. Intell.* **97**(1), 273–324 (1997)
44. Kohavi, R., et al.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *International Joint Conference on Artificial Intelligence* (1995)
45. LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S., Kruschwitz, N.: Big data, analytics and the path from insights to value. In: *MIT Sloan Management Review* (2013)
46. Lee, J., Yoo, S., Choi, S., Kim, H., Hong, C., Seo, J.: Development and application of trim optimization and parametric study using an evaluation system (solution) based on the rans for improvement of eeoI. In: *International Conference on Ocean, Offshore and Arctic Engineering* (2014)
47. Lee, W.S., Bartlett, P.L., Williamson, R.C.: The importance of convexity in learning with squared loss. *IEEE Trans. Inf. Theory* **44**(5), 1974–1980 (1998)
48. Leifsson, L., Saevarsdottir, H., Sigurdsson, S., Vesteinsson, A.: Grey-box modeling of an ocean vessel for operational optimization. *Simul. Model. Pract. Theory* **16**, 923–932 (2008)
49. Lewis, E.V.: *Principles of Naval Architecture*. Society of Naval Architects (1988)
50. Lützen, M., Kristensen, H.: A model for prediction of propulsion power and emissions—tankers and bulk carriers. In: *World Maritime Technology Conference* (2012)
51. MacKay, D.J.C.: *Information Theory, Inference and Learning Algorithms*. Cambridge University Press (2003)
52. Maritime Knowledge Centre: *International shipping facts and figures - Information resources on trade, safety, security, environment*. Technical reports, IMO (2012)
53. Meinshausen, N., Bühlmann, P.: Stability selection. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **72**(4), 417–473 (2010)
54. Moustafa, M.M., Yehia, W., Hussein, A.W.: Energy efficient operation of bulk carriers by trim optimization. In: *International Conference on Ships and Shipping Research* (2015)

55. Ng, A.Y.: Feature selection, l_1 vs. l_2 regularization, and rotational invariance. In: International Conference on Machine Learning (2004)
56. Oneto, L., Ghio, A., Ridella, S., Anguita, D.: Support vector machines and strictly positive definite kernel: the regularization hyperparameter is more important than the kernel hyperparameters. In: International Joint Conference on Neural Networks (2015)
57. Palmé, T., Brehuaus, P., Assadi, M., Klein, A., Kim, M.: New alstom monitoring tools leveraging artificial neural network technologies. In: Turbo Expo: Turbine Technical Conference and Exposition (2011)
58. Petersen, J.P., Winther, O., Jacobsen, D.J.: A machine-learning approach to predict main energy consumption under realistic operational conditions. *Ship Tech. Res.* **59**(1), 64–72 (2012)
59. Quinlan, J.R.: Simplifying decision trees. *Int. J. Man-Mach. Stud.* **27**(3), 221–234 (1987)
60. Reichel, M., Minchev, A., Larsen, N.: Trim optimisation—theory and practice. *Int. J. Marine Navig. Saf. Sea Transp.* **8**(3), 387–392 (2014)
61. Schölkopf, B.: The kernel trick for distances. In: Neural Information Processing Systems (2001)
62. Schölkopf, B., Herbrich, R., Smola, A.J.: A generalized representer theorem. In: Computational Learning Theory (2001)
63. Schultz, M.P., Bendick, J., Holm, E.R., Hertel, W.M.: Economic impact of biofouling on a naval surface ship. *Biofouling* **27**(1), 87–98 (2011)
64. Shao, W., Zhou, P., Thong, S.K.: Development of a novel forward dynamic programming method for weather routing. *J. Mar. Sci. Tech.* **17**(2), 239–251 (2011)
65. Shawe-Taylor, J., Cristianini, N.: Kernel Methods for Pattern Analysis. Cambridge University Press (2004)
66. Simon, N., Friedman, J., Hastie, T., Tibshirani, R.: A sparse-group lasso. *J. Comput. Graph. Stat.* **22**(2), 231–245 (2013)
67. Smith, T.W.P., Jalkanen, J.P., Anderson, B.A., Corbett, J.J., Faber, J., Hanayama, S., OKeffe, E., Parker, S., Johansson, L., Aldous, L.: Third imo ghg study 2014. Technical report, International Maritime Organisation (2014)
68. Stewart, T.R., McMillan Jr, C.: Descriptive and prescriptive models for judgment and decision making: implications for knowledge engineering. In: Expert Judgment and Expert Systems (1987)
69. Stopford, M.: Maritime Economics. Routledge, New York (2009)
70. Sugumaran, V., Muralidharan, V., Ramachandran, K.: Feature selection using decision tree and classification through proximal support vector machine for fault diagnostics of roller bearing. *Mech. Syst. Signal Process.* **21**(2), 930–942 (2007)
71. Tibshirani, R.: Regression shrinkage and selection via the lasso. *J. Royal Stat. Soc. Ser. B (Methodol.)* pp. 267–288 (1996)
72. Tikhonov, A., Arsenin, V.Y.: Methods for solving ill-posed problems. Nauka, Moscow (1979)
73. UNCTAD: Review of maritime transport. Technical report, United Conference on Trade and Development (2012)
74. Vapnik, V.N.: Statistical Learning Theory. Wiley–Interscience (1998)
75. Von Karman, T., Gabrielli, G.: What price speed? specific power required for propulsion of vehicles. *Mech. Eng.* **72**, 775–781 (1950)
76. Wang, H., Faber, J., Nelissen, D., Russell, B., St Amand, D.: Marginal Abatement Costs and Cost Effectiveness of Energy-Efficiency Measures. Technical report, Institute of Marine Engineering, Science and Technology (2010)
77. White, A.P., Liu, W.Z.: Technical note: bias in information-based measures in decision tree induction. *Mach. Learn.* **15**(3), 321–329 (1994)
78. Widodo, A., Yang, B.S.: Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **21**(6), 2560–2574 (2007)
79. Yoon, H., Yang, K., Shahabi, C.: Feature subset selection and feature ranking for multivariate time series. *IEEE Trans. Knowl. Data Eng.* **17**(9), 1186–1198 (2005)
80. Young, D.M.: Iterative Solution of Large Linear Systems. Dover Publications, Com (2003)
81. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. *J. R. Stat. Soc.: Ser. B (Stat. Method.)* **67**(2), 301–320 (2005)
82. Zou, H., Hastie, T., Tibshirani, R.: On the degrees of freedom of the lasso. *Ann. Stat.* **35**(5), 2173–2192 (2007)