Hybrid Modelling Approach of a Four-stroke Medium Speed Diesel Engine

Coraddu, A.; Kalikatzarakis, Miltos; Geertsma, R.D.; Oneto, Luca

Publication date
2021

Document Version
Final published version

Published in
Proceedings of the 13th Symposium on High-Performance Marine Vehicles, HIPER '21

Hybrid Modelling Approach of a Four-stroke Medium Speed Diesel Engine

Andrea Coraddu, University of Strathclyde, Glasgow/UK, andrea.coraddu@strath.ac.uk
Miltos Kalikatzarakis, University of Strathclyde, Glasgow/UK, miltos.kalikatzarakis@strath.ac.uk
Rinze Geertsma, Delft University of Technology, Delft/The Netherlands, R.D.Geertsma@tudelft.nl
Luca Oneto, Università Degli Studi di Genova, Genova/Italy, luca.oneto@unige.it

Abstract

Diesel engines will remain a fundamental component of propulsion systems due to their maturity, reliability, and power density. Building Digital Twins of the propulsion system is one feasible solution to pursue the optimal propulsions system operation, estimating system states and efficiency. This work will investigate a modelling approach that combines high accuracy while satisfying real-time prediction capabilities by coupling a physics-based model with a data-driven modelling approach. We will demonstrate that the proposed hybridisation framework can provide state-of-the-art prediction capabilities in real-time, utilising operational data from a turbocharged, four-stroke medium-speed diesel engine.

1. Introduction

In recent years the maritime industry is confronted by several challenges, including volatile bunker prices, *García-Martos et al. (2013)* that affect cargo transportation costs and the shipowners' competitiveness and viability of their operations, and strict regulations to limit emissions and their environmental impact to reduce CO_2 emissions from shipping by 40-50%, *European Commission (2013a)*. As a result of this combination, the issue of energy efficiency and environmental sustainability of maritime operations is currently prioritised in the maritime industry, with shipowners and operators adopting measures to lower fuel consumption and associated emissions, and researchers studying innovative technologies and methods that can increase the environmental efficiency and cost-effectiveness of ship operations.

Propulsion system design is facing the challenge of continuously rising complexity to satisfy these demands. However, energy efficiency should not only be a design issue but also be preserved in operation. In pursuing the optimal propulsion system operation, estimating system states and efficiency is of great importance. In this respect, building a Digital Twin (DT) of the propulsion system that coexists during operation, providing predictions and offering insight into the operation, is one feasible solution. A critical requirement for the DT is the need for a modelling approach that can precisely reflect the characteristics intrinsic to the propulsion plant and precisely predict the state of its physical counterpart under all operating conditions in real-time, as reported in *Bondarenko and Fukuda (2020).* Main engines are the main factors of energy loss and emission production on-board and will remain an unavoidable part of propulsion systems due to their maturity, reliability, and power density, *Baldi et al. (2014, 2015).* Diesel engine (DE) modelling has evolved over the years, and various types of models can be found in the literature, with varying degrees of computational complexity and prediction quality. Most widely employed are Mean Value Engine Models (MVEMs), *Guan et al. (2014), Grimmelius et al. (2010); Geertsma et al. (2017),* which provide adequate accuracy in the prediction of most engine parameters while being computationally cheap, and zero-dimensional (0D) models that operate on per per-crank basis, allowing the calculation of parameters of the gas within the engine cylinders as reported in *Sapra et al. (2020), Catania et al. (2011), Asad et al. (2014*).

Suppose the requirements for a modelling approach include real-time prediction of the main engine performance parameters with a high degree of accuracy. In that case, neither MVEM nor 0D models are applicable due to moderate prediction capabilities (MVEMs) and computational time requirements (0D models). More sophisticated approaches, with respect to MVEMs and 0D, are one-dimensional (1D) and three-dimensional (3D) models that operate on a per-crank basis, *Merker et al. (2005).* These approaches are more computationally demanding compared to MVEMs. However, they can predict the

detailed gas processes inside the cylinders with higher accuracy, *Mohammadkhani et al. (2019)*. Several attempts to combine MVEM and 0D, 1D, or 3D models have been proposed, enhancing the predictive abilities of MVEMs with lower computational requirements than their 0D, 1D, or 3D counterparts, as suggested by *Livanos et al. (2007)* and *Ding et al. (2010)*. For instance, *Baldi et al. (2015)* combined MVEM and 0D models to investigate the propulsion behaviour of a Handymax-size product carrier under constant and variable engine speed operations. PMs can adequately capture most process parameters of a DE under a broad range of operating conditions. However, there is a clear trade-off between accuracy and computational requirements. The most accurate 3D models cannot run in real-time, whereas MVEMs lack accuracy, especially during transient operations.

DDMs have been successfully applied in a variety of maritime applications, provided that the necessary quality and quantity of historical data is available, as reported in *Coraddu et al. (2017,2019a,2019b, 2020,2021a,2021b)* and *Cipollini et al. (2018a, 2018b)*. For instance, *Nikzadfar and Shamekhi (2014)* developed an Artificial Neural Network (ANN) to study the relative contribution of several operating parameters to the performance of a DE. The ability of ANNs to predict performance parameters of a DE was also demonstrated in *Özener et al. (2013)* to predict a variety of performance parameters and emissions. A hydrogen dual-engine for automotive applications was the case study of *Syed et al. (2017)*: ANNs proved to be highly efficient to predict specific fuel consumption and a variety of emissions.

HMs are a pretty recent modelling approach, especially in the maritime field, and just very few works showed the advantages of a hybrid approach regarding pure PMs and DDMs, as reported in *Coraddu et al. (2018, 2021a)* and *Miglianti et al. (2019, 2020).* For instance, *Coraddu et al. (2017)* show that it is possible to predict fuel consumption with HMs effectively. Moreover, *Coraddu et al.(2018, 2021a)* attempted to model the engine exhaust gas temperature with HMs under steady-state and transient conditions. *Mishra and Subbarao (2021)* compared the performance of a PM, a DDM, and an HM to predict dynamic combustion control parameters of a Reactivity Controlled Compression Ignition engine across five engine loads. The parameters included the start of combustion, the 50% mass fraction burnt crank angle, and combustion peak pressure. The authors compared the model predictions with measured data from experiments, concluding that the prediction capability of the HM was far superior to the DDM and PM across all parameters. *Bidarvatan et al. (2014)* developed an HM to predict several performance parameters of Homogeneous Charge Compression Ignition (HCCI) engines. Namely, the 50% mass fraction burnt crank angle, the indicated mean effective pressure, exhaust temperature, and concentration of CO, total unburned hydrocarbons and NOx. The proposed HM combined a PM and 3 ANNs, designed to minimise computational time requirements. The authors compared the predictions of the proposed HM with experimental data at 309 steady-state and transient conditions for two HCCI engines concluding that the HM offered approximately 80% better accuracy compared to the PM, or 60% compared to the DDM.

The amount of literature available on the HMs is limited, as this is a relatively new research field. Moreover, focusing on the marine DE applications, a consistent and clear description of a modelling framework for marine DEs able to hybridise PMs and DDMs is not yet readily available.

This work will investigate a modelling approach that combines high accuracy whilst satisfying real-time prediction capabilities by coupling a physics-based, low-computational MVEM with a Data-Driven model. We will demonstrate that combining these two approaches in a hybrid modelling framework can provide state-of-the-art prediction capabilities in real-time, utilising several months of operational data from a turbocharged, four-stroke medium-speed DE. With this in mind, first, a 0D DE model, already available in *Kalikatzarakis et al. (2021)*, is briefly described. Subsequently, different DDMs will be developed, tested, and compared. These models will leverage the information encapsulated in historical data to produce accurate predictions on a set of performance parameters of the DE. Finally, we will present the hybridisation framework where HM will be proposed, leveraging on both the DDM and the PM previously developed. The authors will showcase the performance in terms of accuracy, reliability, and computational requirements of the HM, demonstrating the superiority of the proposed hybridisation framework on a comprehensive dataset containing operational data from a marine DE for a time of approximately three years.

2. Physical Models

In this section, we report an overview of the DE modelling approach for the sake of completeness. A more detailed explanation and the validation results are available in *Kalikatzarakis et al. (2021)*. The vessel's DEs have been modelled utilising a modular approach. Inputs include the geometric data of the engine, the intake and exhaust valves profiles, the compressor and turbine performance maps, the waste gate geometric and control details, the constants of engine sub-models (combustion, heat transfer and friction), the engine operating point (load/speed), and the ambient conditions. Initial conditions are required for the temperature, pressure and composition of the working medium contained in the engine cylinders, pipes, and receivers. The engine scavenging air and exhaust gas receivers are modelled as flow receiver elements (control volumes), whereas flow elements represent the compressor, air cooler, cylinders, and turbine. The engine boundaries are modelled using fixed fluid elements of constant pressure and temperature, and shaft elements are utilised to compute the rotational speed of the turbocharger and crankshaft. The governor of the engine is responsible for adjusting the fuel rack position and incorporates the appropriate fuel rack limiters. Finally, air and exhaust gas properties are dependent on temperature, fuel-air equivalence ratio and composition. For the calculation of the exhaust gas composition, oxygen, nitrogen, carbon dioxide and steam were considered. All flow elements are modelled using the open thermodynamic system concept of *Watson and Janota (1982)* and *Heywood (1988)*, and use as inputs the pressure (p), temperature (T), and the properties of the working medium contained in the adjacent elements. Subsequently, mass ($\dot{m}$) and energy flow rates entering and exiting each element are computed by applying the mass and energy conservation laws. These flow rates are further provided as inputs in the adjacent flow receiver elements, whereas torque outputs are utilised as inputs in the shaft elements. The latter, through the angular momentum conservation, compute the rotational speeds of the turbocharger (ω_{tc}) and propulsion plant shafts. The compressor is modelled using its steady-state performance map, estimated utilising the method proposed by *Casey and Robinson (2013)*, whereas the turbine is modelled using its swallowing capacity and efficiency map. Moreover, the pressures losses occurring in the air cooler and air filter are dependent on the air mass flow rate, as is the air cooler effectiveness. No heat transfer is considered in the model of the scavenging air receiver, whereas the transferred heat to the ambient from the gas in the exhaust gas receiver is calculated from the overall heat transfer coefficient, using a Nusselt-Reynolds number correlation for gas flowing in the pipes according to *Rohsenow and Hartnett (1988)*. Moreover, pressure losses in the exhaust gas receiver are dependent on the exhaust gas mass flow rate. The in-cylinder process is described by a two-zone zero-dimensional model as in *Merker et al. (2005)*. This type of model operates on per crank-angle basis, using the mass and energy conservation equations, along with the gas state equation, which are solved in their differential form, so that the parameters of the gas within the engine cylinders and manifolds, such as pressure, temperature and gas composition can be calculated. Combustion is modelled through a two-zone model, considering a zone containing the combustion products and an unburned mixture zone according to *Merker et al. (2005)*. The Woschni heat transfer model, originating from *Woschni (1967)*, and employed extensively in various studies, is utilised to compute the in-cylinder gas to wall heat transfer coefficient, *Merker et al. (2005)*. According to the Vibe model, the heat release rate is simulated, as described in *Merker et al. (2005)*. The combustion products are evaluated utilising the method of *Rakopoulos et al. (1994)*, due to its minimal computational time requirements and reasonable agreement with experiments: for the burning zone, given its volume, temperature, mass of fuel burnt and mass of air entrained, the concentration of each gas composition species discussed in *Kalikatzarakis et al. (2021)*, can be evaluated by solving an 11 × 11 non-linear system obtained from 7 non-linear equilibrium equations and 4 linear atom balance equations. Thermal NO has been evaluated according to the extended Zeldovich mechanism, for which the reaction rates were selected according to *Hanson and Salimian (1984)*.

3. Hybrid Models

This section will introduce our hybridisation framework, starting from a formal description of the DDMs.

3.1. Data-Driven Models

To develop a fast yet accurate dynamic model of a four-stroke marine DE, a general modelisation framework is here defined, characterised by an input space $\boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^d$, an output space $\boldsymbol{\mathcal{Y}} \subseteq \mathbb{R}^b$and an unknown relation $\mu: \boldsymbol{\mathcal{X}} \rightarrow \boldsymbol{\mathcal{Y}}$ to be learned, *Shalev-Shwartz and Ben-David (2014)*, *Hamilton (2020)*. In this work, $\boldsymbol{\mathcal{X}}$ is composed by the measurements available from the engine monitoring system (see Section 4), while the output space $\boldsymbol{\mathcal{Y}}$ refers to the target features accounting for the engine fuel consumption $\dot{m}_f$, turbocharger rotational speed N_{tc}, turbine outlet temperature $T_{\mathrm{t,out}}$ and exhaust manifold temperature T_{er}. We define the model $h: \boldsymbol{\mathcal{X}} \rightarrow \boldsymbol{\mathcal{Y}}$ as an artificial simplification of μ. The model h, as described in Section 1, can be obtained with different kinds of techniques, for example, requiring some physical knowledge of the problem, as in PMs (see Section 2), or the acquisition of large amounts of data, as in DDMs or using both information (see Section 3). Between the DDMs, it is possible to identify two families of approaches, *Shalev-Shwartz and Ben-David (2014).* The first one, comprising traditional Machine Learning (ML) methods, needs an initial phase where the features must be defined a-priori from the data via feature engineering or implicit or explicit feature mapping, Shawe-Taylor and *Cristianini (2004).* The second family, which includes Deep Learning (DL) methods, automatically learns both the features and models from the data, *Goodfellow et al. (2016).* For small cardinality datasets and outside particular applications (e.g., computer vision and natural language processing), DL does not perform well since they require a large amount of data to be reliable and to outperform traditional ML models, *Wainberg et al. (2016).* In the ML, the above-mentioned can be easily mapped in a typical regression problem, *Vapnik (1998).* In fact, ML techniques aim at estimating the unknown relationship μ between input and output through a learning algorithm $\mathcal{A}_{\mathcal{H}}$ which exploits some historical data to learn h and where $\mathcal{H}$ is a set of hyperparameters that characterises the generalisation performance of $\mathcal{A}$, Oneto (2020). The historical data consists of a series of n examples of the input/output relation μ and are defined as $\mathcal{D}_n = \{(\boldsymbol{x}_1, y_1), \cdots, (\boldsymbol{x}_n, y_n)\}$ where $x \in \boldsymbol{\mathcal{X}}$ and $y \in \boldsymbol{\mathcal{Y}}$.

This paper will leverage ML models from the Kernel Methods family called Kernel Regularised Least Squares (KRLS), *Vovk (2013).* The idea behind KRLS can be summarised as follows. During the training phase, the quality of the learned function $h(x)$ is measured according to a loss function $\ell(h(\boldsymbol{x}), y)$, as reported in *Rosasco et al. (2004),* with the empirical error

$$\hat{L}_n(h) = \frac{1}{n}\sum_{i=1}^{n} \ell(h(\boldsymbol{x}_i), y_i). \tag{1}$$

A simple criterion for selecting the final model during the training phase could then consist of simply choosing the approximating function that minimises the empirical error $\hat{L}_n(h)$. This approach is known as Empirical Risk Minimisation (ERM), *Vapnik (1998).* However, ERM is usually avoided in ML as it leads to severe overfitting of the model on the training dataset. As a matter of fact, in this case, the training process could choose a model complicated enough to perfectly describe all training samples (including the noise, which afflicts them). In other words, ERM implies memorisation of data rather than learning from them. A more effective approach is to minimise a cost function where the trade-of between accuracy on the training data and a measure of the complexity of the selected model is achieved, *Tikhonov and Arsenin (1979),* implementing the Occam's razor principle

$$h^*: \min_h \hat{L}_n(h) + \lambda\, \mathrm{C}(h). \tag{2}$$

The best-approximating function h^* is chosen as the one that is complicated enough to learn from data without overfitting them, where C(·) is a complexity measure: depending on the exploited ML approach, different measures are realised. $\lambda \in [0, \infty)$ is a hyperparameter, that must be set a-priori and is not obtained as an output of the optimisation procedure: it 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 C(·). The optimal value for λ is problem-dependent, and tuning this hyperparameter is a non-trivial task, as will be discussed later in this section. In KRLS, models are defined as

$$h(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{\varphi}(\boldsymbol{x}), \tag{3}$$

where $\boldsymbol{\varphi}$ is an a-priori defined Feature Mapping (FM) as reported in *Shalev-Shwartz and Ben-David (2014)*, allowing to keep the structure of $h(\boldsymbol{x})$ linear.

The complexity of the models in KRLS is measured as

$$\mathrm{C}(h) = \| \boldsymbol{w} \|^2, \tag{4}$$

i.e., the Euclidean norm of the set of weights describing the regressor, which is a standard complexity measure in ML, *Vovk (2013)*. Regarding the loss function, the square loss is typically adopted because of its convexity, smoothness, and statistical properties, as shown in *Rosasco et al. (2004)*

$$\hat{L}_n(h) = \frac{1}{n}\sum_{i=1}^{n} \ell(h(\boldsymbol{x}_i), y_i) = \frac{1}{n}\sum_{i=1}^{n} [h(\boldsymbol{x}_i) - y_i]^2 \tag{5}$$

Consequently, problem (2) can be reformulated as

$$\boldsymbol{w}^*: \min_{\boldsymbol{w}} \sum_{i=1}^{n} [\boldsymbol{w}^T \boldsymbol{\varphi}(\boldsymbol{x}) - y_i]^2 + \lambda \| \boldsymbol{w} \|^2. \tag{6}$$

By exploiting the Representer Theorem, *Schölkopf et al. (2001)*, the solution h^* of the problem (6) can be expressed as a linear combination of the samples projected in the space defined by $\boldsymbol{\varphi}$

$$h^*(\boldsymbol{x}) = \sum_{i=1}^{n} \iota_i \boldsymbol{\varphi}(\boldsymbol{x}_i)^T \boldsymbol{\varphi}(\boldsymbol{x}). \tag{7}$$

It is worth underlining that, according to the kernel trick, it is possible to reformulate $h^*(\boldsymbol{x})$ without explicit knowledge of $\boldsymbol{\varphi}$, and consequently avoiding the curse of dimensionality of computing $\boldsymbol{\varphi}$, by using a proper kernel function $K(\boldsymbol{x}_i, \boldsymbol{x}) = \boldsymbol{\varphi}(\boldsymbol{x}_i)^T \boldsymbol{\varphi}(\boldsymbol{x})$

$$h^*(\boldsymbol{x}) = \sum_{i=1}^{n} \iota_i K(\boldsymbol{x}_i, \boldsymbol{x}). \tag{8}$$

Several kernel functions can be retrieved in literature, such as those reported in *Schölkopf (2001)* and *Cristianini and Shawe-Taylor (2000)*, each with a particular property that can be exploited the problem under exam. Usually, the Gaussian kernel is chosen

$$K(\boldsymbol{x}_i, \boldsymbol{x}) = e^{-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}\|^2}, \tag{9}$$

because of the theoretical reasons described in *Keerthi and Lin (2003)* and *Oneto et al. (2015)*, and because of its effectiveness as reported in *Fernández-Delgado et al. (2014)* and *Wainberg et al. (2016)*. γ is another hyperparameter that regulates the solution's nonlinearity that must be tuned, as explained later. The Gaussian kernel can implicitly create an infinite dimensional $\boldsymbol{\varphi}$ and thanks to this, the KRLS can learn any possible function, *Keerthi and Lin (2003)*. The KRLS problem of Eq. (6) can be reformulated by exploiting kernels as

$$\boldsymbol{\iota}^*: \min_{\boldsymbol{\iota}} \| Q\boldsymbol{\iota} - y \|^2 + \lambda \boldsymbol{\iota}^T Q \boldsymbol{\iota}, \tag{10}$$

where $y = [y_1, \dots, y_n]^T$, $\boldsymbol{\iota} = [\iota_1, \dots, \iota_n]^T$, the matrix Q such that $Q_{i,j} = K(\boldsymbol{x}_j, \boldsymbol{x}_i)$, and the identity matrix $I \in \mathbb{R}^{n \times n}$. By setting the gradient equal to zero w.r.t. $\boldsymbol{\iota}$ it is possible to state that

$$(Q + \lambda I)\boldsymbol{\iota}^* = y, \tag{11}$$

which is a linear system for which effective solvers have been developed over the years, allowing it to cope with even very large sets of training data, *Young (2003)*.

3.2. Model Selection and Error Estimation

The problems that still must be faced are how to tune the hyperparameters (λ and γ) and estimate the performance of the final model. Model Selection (MS) and Error Estimation (EE) deal precisely with these problems, *Oneto (2020)*. Researchers and practitioners commonly use resampling techniques since they work well in most situations, and therefore we will exploit them in this work. Other alternatives exist based on the Statistical Learning Theory, but they tend to underperform resampling techniques in practice, as demonstrated by *Oneto (2020)*. Resampling techniques are based on a simple

idea: the original dataset $\mathcal{D}_n$ is resampled once or many (n_r) times, with or without replacement, to build three independent datasets called learning, validation and test sets, respectively $\mathcal{L}_l^r$, $\mathcal{V}_v^r$, and $\mathcal{T}_r^r$, with $r \in \{1, \cdots, n_r\}$ such that

$$\mathcal{L}_l^r \cap \mathcal{V}_v^r = \emptyset, \mathcal{L}_l^r \cap \mathcal{T}_t^r = \emptyset, \mathcal{V}_v^r \cap \mathcal{T}_t^r = \emptyset, \mathcal{L}_l^r \cup \mathcal{V}_v^r \cup \mathcal{T}_t^r = \mathcal{D}_n. \quad (12)$$

Subsequently, to select the best hyperparameters' combination $\mathcal{H} = \{\lambda, \gamma\}$ considering a set of possible ones $\boldsymbol{\mathcal{H}} = \{\mathcal{H}_1, \mathcal{H}_2, \cdots\}$ for the algorithm $\mathcal{A}_{\mathcal{H}}$ or, in other words, to perform the MS phase, the following procedure must be applied:

$$\mathcal{H}^*: \arg\min_{\mathcal{H} \in \boldsymbol{\mathcal{H}}} \sum_{r=1}^{n_r} \mathrm{M}(\mathcal{A}_{\mathcal{H}}(\mathcal{L}_l^r), \mathcal{V}_v^r), \quad (13)$$

where $h = \mathcal{A}_{\mathcal{H}}(\mathcal{L}_l^r)$ is a model built with the algorithm $\mathcal{A}$ with its set of hyperparameters $\mathcal{H}$ and with the data $\mathcal{L}_l^r$, and where $M(h, \mathcal{V}_v^r)$ is a desired metric. Since the data in $\mathcal{L}_l^r$ is independent of the data in $\mathcal{V}_v^r$, $\mathcal{H}^*$ should be the set of hyperparameters that allows achieving a small error on a data set that is independent of the training set. Then, to evaluate the performance of the optimal model, which is $h_{\mathcal{A}}^* = \mathcal{A}_{\mathcal{H}^*}(\mathcal{D}_n)$ or, in other words, to perform the EE phase, the following procedure must be applied:

$$\mathrm{M}(h_{\mathcal{A}}^*) = \frac{1}{n_r} \sum_{r=1}^{n_r} \mathrm{M}(\mathcal{A}_{\mathcal{H}^*}(\mathcal{L}_l^r \cup \mathcal{V}_v^r), \mathcal{T}_t^r) \quad (14)$$

Since the data in $\mathcal{L}_l^r \cap \mathcal{V}_v^r$ are independent from the ones in $\mathcal{T}_t^r$, $\mathrm{M}(h_{\mathcal{A}}^*)$ is an unbiased estimator of the true performance, measured with the metric M, of the final model, *Oneto (2020).* In this work, we will rely on Complete k-fold cross validation, which means setting

$$n_r \leq \binom{n}{k}\binom{n - \frac{n}{k}}{k}, l = (k-2)\frac{n}{k}, v = t = \frac{n}{k}, \quad (14)$$

and resampling without replacement. Note that, in our application, we have a further constraint in terms of dependence in time between the samples. For this reason, when resampling the data form $\mathcal{D}_n$we keep data of different periods in $\mathcal{L}_l^r$, $\mathcal{V}_v^r$, and $\mathcal{T}_t^r$As reported in *Hamilton (2020).* For what concerns the applied metric M, we will rely on the Mean Absolute Error (MAE), the Mean Absolute Percentage of Error (MAPE), and the Pearson Product-Moment Correlation Coefficient PPMCC according to *Willmott and Matsuura (2005).* Since in regression, it is pretty hard to synthesise the quality of a predictor in a single metric, we will also rely on visualisation techniques like the scatter plot and histograms, *Shao et al. (2017).*

3.2. Hybrid Models

In this section, we depict a framework able to consider both the physical knowledge about the problem encapsulated in the PMs of Section 2 and the information hidden in the available data as the DDMs of Section 3.1. For this purpose, we will start from a simple observation: an HM, based on the previous observation, should learn from the data without being too different or too far away from the PM. From the Data Science and ML point of view, this requirement can be straightforwardly mapped to a typical ML Multi Task Learning (MTL) problem, *Baxter (2000), Caruana (1997), Evgeniou and Pontil (2004), Bakker and Heskes (2003), Argyriou et al. (2008).* MTL aims at simultaneously learning two concepts, in this case the PM and the available data, through a learning algorithm $\mathcal{A}_{\mathcal{H}}$ which exploits the data in $\mathcal{D}_n$ to learn a function h which is both close to the observation, the data $\mathcal{D}_n$ and the PM, namely its forecasts. Consequently, in this case, a slightly different scenario is presented where the dataset is composed of a triple of points $\mathcal{D}_n$= $\{(\boldsymbol{x}_1, y_1, p_1), \cdots, (\boldsymbol{x}_n, y_n, p_n)\}$ where p_i is the output of the PM in the point $\boldsymbol{x}_n$with $i \in \{1, \cdots, \mathrm{N}\}$. The target is to learn a function able to approximate both μ, namely the relation between the input $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ and the output $y \in \boldsymbol{\mathcal{Y}}$, and the PM, namely, the relation between the input and the output of the PM. Two tasks must be learned, and for this purpose, there are two main approaches: the first approach is called Shared Task Learning (STL) and the second Independent Task Learning (ITL). While the latter independently learn a different model for each task, the former aims to learn a model that is common between all tasks. A well-known weakness of these methods is that they tend to generalise poorly on one of the two tasks, *Baxter (2000).* In this work, we show that an appealing

approach to overcome such limitations is provided by MTL as suggested by *Baxter (2000)* and *Argyriou et al. (2008).* This methodology leverages the information between the tasks to learn more accurate models. To apply the MTL approach to this case, it is possible to modify the KRLS problem of Eq. (6) to simultaneously learn a shared model and a task specific model which should be close to the shared model. In this way, we obtain a model which can simultaneously learn the two tasks. The model we are interested in is the shared model, while the task specific models are just used as a tool. A shared model is defined as $h(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x})$, and two task specific models as

$$h_i(\boldsymbol{x}) = \boldsymbol{w}_i^T\boldsymbol{\varphi}(\boldsymbol{x}), \quad i \in \{y, p\}. \tag{15}$$

Then, it is possible to state the MTL version of Eq. (6), as follows

$$\begin{aligned}\boldsymbol{w}^*, \boldsymbol{w}_y^*, \boldsymbol{w}_p^*: \min_{\boldsymbol{w},\boldsymbol{w}_y,\boldsymbol{w}_p} &\sum_{i=1}^{n} [\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}) - y_i]^2 + [\boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}) - p_i]^2 \\ &+ \sum_{i=1}^{n} \left[\boldsymbol{w}_y^T\boldsymbol{\varphi}(\boldsymbol{x}) - y_i\right]^2 + \left[\boldsymbol{w}_p^T\boldsymbol{\varphi}(\boldsymbol{x}) - p_i\right]^2 \\ &+\lambda \| \boldsymbol{w} \|^2 + \varkappa\left(\|\boldsymbol{w} - \boldsymbol{w}_y\|^2 + \|\boldsymbol{w} - \boldsymbol{w}_p\|^2\right)\end{aligned} \tag{16}$$

where λ is the usual regularisation of KRLS and $\varkappa \in [0, \infty)$, instead, is another hyperparameter that forces the shared model to be close to the task specific models. Basically, the MTL problem of Eq. (16) is a concatenation of three learning problems solved with KRLS plus a term that tries to keep a relation between all the three different problems. By exploiting the kernel trick as in KRLS, it is possible to reformulate the problem (16), as follows

$$\boldsymbol{\iota}^*: \min_{\iota} \left\| \begin{bmatrix} Q & Q & 0 & 0 \\ Q & Q & 0 & 0 \\ 0 & 0 & Q & 0 \\ 0 & 0 & 0 & Q \end{bmatrix} \boldsymbol{\iota} - \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{p} \\ \boldsymbol{y} \\ \boldsymbol{p} \end{bmatrix} \right\|^2 + \boldsymbol{\iota}^T \begin{bmatrix} (\lambda + 2\varkappa)Q & (\lambda + 2\varkappa)Q & -\varkappa Q & -\varkappa Q \\ (\lambda + 2\varkappa)Q & (\lambda + 2\varkappa)Q & -\varkappa Q & -\varkappa Q \\ -\varkappa Q & -\varkappa Q & \varkappa Q & 0 \\ -\varkappa Q & -\varkappa Q & 0 & \varkappa Q \end{bmatrix} \boldsymbol{\iota} \tag{17}$$

where $\boldsymbol{p} = [p_1, \dots, yp_n]^T$.

The solution of this problem is again equivalent to solving a linear system

$$\begin{bmatrix} Q + (\lambda + 2\varkappa)I & Q + (\lambda + 2\varkappa)I & -xI & -\varkappa I \\ Q + (\lambda + 2\varkappa)I & Q + (\lambda + 2\varkappa)I & -xI & -\varkappa I \\ -\varkappa I & -\varkappa I & Q + \varkappa I & 0 \\ -\varkappa I & -\varkappa I & 0 & Q + \varkappa I \end{bmatrix} \boldsymbol{\iota}^* = \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{p} \\ \boldsymbol{y} \\ \boldsymbol{p} \end{bmatrix} \tag{18}$$

The function that the authors are interested in, the shared one, can be expressed as follows

$$h(\boldsymbol{x}) = \boldsymbol{w}^T\boldsymbol{\varphi}(\boldsymbol{x}) = \sum_{i=1}^{n} (\iota_i + \iota_{i+n})K(\boldsymbol{x}_i, \boldsymbol{x}). \tag{19}$$

4. Data Description

Data from a naval vessel equipped with a MAN B&W V28-33D medium speed four-stroke DE, Table I, has been exploited in this work.

Table I: Main characteristics of the MAN 12 V28-33D engine

Feature	Value	Unit	Feature	Value	Unit
Cylinders	V12, 16, 20	[-]	Brake power at 60% MCR	3240	[kW]
Bore diameter	280	[mm]	Brake power at 80% MCR	4320	[kW]
Stroke length	330	[mm]	Brake power at MCR	5400	[kW]
Number of cylinders	12	[-]	Mean Effective Pressure	26.9	[bar]
Revolutions per cycle	2	[-]	Mean Piston Speed	11	[m/s]
Engine speed at MCR	1000	[rpm]	Specific Fuel consumption	191	[g/kWh]

The DE is installed on board one of the Holland Class Oceangoing Patrol Vessels. The propulsion system of the vessel consists of two shafts with Controllable Pitch Propellers (CPP), a gearbox, and one DE per shaft. The vessel is equipped with a data logging system used both for on board monitoring and control and for land-based performance analysis. The dataset utilised consists of two different data sources, Table II: standard measurements (steady-state) performed during Shop Trials (ST) that were used to calibrate the PM model (see Section 2), *Kalikatzarakis et al. (2021)*, and operational data originating from the vessel's data logging system, used by the ship operator for performance monitoring purposes, which has been exploited to evaluate the performance of the PM model in dynamic conditions (see Section 2), and training, validate, and test the DDMs and HMs (see Sections 3.1 and 3).

Table II: Table captions above table

Variable Name	Symbol	Unit
Timestamp	t	[hh:mm:ss]
Governor Position	G_p	[-]
Engine Rotational Speed	n_e	[rpm]
Engine Torque	M_e	[kNm]
Charge Air Temperature at Scavenging Receiver	T_{sc}	[°C]
Charge Air Temperature at Compressor Inlet	$T_{\mathrm{c,in}}$	[°C]
Charge Air Temperature at Compressor Outlet	$T_{\mathrm{c,out}}$	[°C]
Exhaust Gas Temperature at Turbine Outlet	$T_{t,\mathrm{out}}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},1}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},2}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},3}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},4}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},5}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},6}$	[°C]
Main Bearing Temperature	$T_{\mathrm{b},7}$	[°C]
Lube Oil Compartment No. 1 Temperature	$T_{\mathrm{l},1}$	[°C]
Lube Oil Compartment No. 2 Temperature	$T_{\mathrm{l},2}$	[°C]
Lube Oil Compartment No. 3 Temperature	$T_{l,3}$	[°C]
Lube Oil Compartment No. 4 Lube Oil	$T_{\mathrm{l},4}$	[°C]
Lube Oil Compartment No. 5 Lube Oil	$T_{l,5}$	[°C]
Lube Oil Engine Inlet Temperature	$T_{\mathrm{le,in}}$	[°C]
Lube Oil Engine Outlet Temperature	$T_{\mathrm{le,out}}$	[°C]
High-Temperature Sea Cooling Water - Inlet	$T_{\mathrm{ht,in}}$	[°C]
High-Temperature Sea Cooling Water - Outlet	$T_{\mathrm{ht,out}}$	[°C]
Low-Temperature Sea Cooling Water - Inlet	$T_{\mathrm{lt,in}}$	[°C]
Low-Temperature Sea Cooling Water - Outlet	$T_{\mathrm{lt,out}}$	[°C]
Fuel Oil Supply Temperature	T_f	[°C]
Charge Air Temperature at Compressor Outlet – Bank A	$T_{\mathrm{c,out}}^{\mathrm{A}}$	[°C]
Charge Air Temperature at Compressor Outlet – Bank B	$T_{\mathrm{c,out}}^{\mathrm{B}}$	[°C]
Charge Air Temperature at Compressor Inlet – Bank A	$T_{\mathrm{c,in}}^{\mathrm{A}}$	[°C]
Charge Air Temperature at Compressor Outlet – Bank B	$T_{\mathrm{c,out}}^{\mathrm{B}}$	[°C]
Charge Air Engine Inlet Pressure	$p_{\mathrm{ca,in}}$	[Pa]
Charge Air Engine Inlet Temperature	$T_{\mathrm{ca,in}}$	[°C]
Fuel Consumption	$\dot{m}_f$	[kg/h]
TC rotational speed	N_{tc}	[rpm]
Turbine Outlet Temperature	$T_{\mathrm{t,out}}$	[°C]
Exhaust Receiver Temperature	T_{er}	[°C]

5. Experimental Results

In this section, we test the models developed in Sections 2 and Sections 3 on the data described in Section 4, comparing the performance of PMs, DDMs, and HMs in dynamic operational conditions. First, we report the hyperparameters ranges for the DDM and HM.

For the DDM, the set of hyperparameters tuned during the MS phase are $\mathcal{H} = \{\lambda, \gamma\}$ chosen in $\boldsymbol{\mathcal{h}} = \{10^{-4.0}, 10^{-3.8}, \cdots, 10^{+4.0}\} \times \{10^{-4.0}, 10^{-3.8}, \cdots, 10^{+4.0}\}$. For the HM, the set of hyperparameters tuned during the MS phase are $\mathcal{H} = \{\lambda, \gamma, \varkappa\}$ chosen in $\boldsymbol{\mathcal{h}} = \{10^{-4.0}, 10^{-3.8}, \cdots, 10^{+4.0}\} \times \{10^{-4.0}, 10^{-3.8}, \cdots, 10^{+4.0}\} \times \{10^{-4.0}, 10^{-3.8}, \cdots, 10^{+4.0}\}$.

All the tests have been repeated 30 times, and the average results are reported together with their t-student 95% confidence interval to ensure the statistical validity of the results. Table III reports the performance (MAE, MAPE, and PPMCC) of the different models (PM, DDM, and HM) for the different targets using to predict.

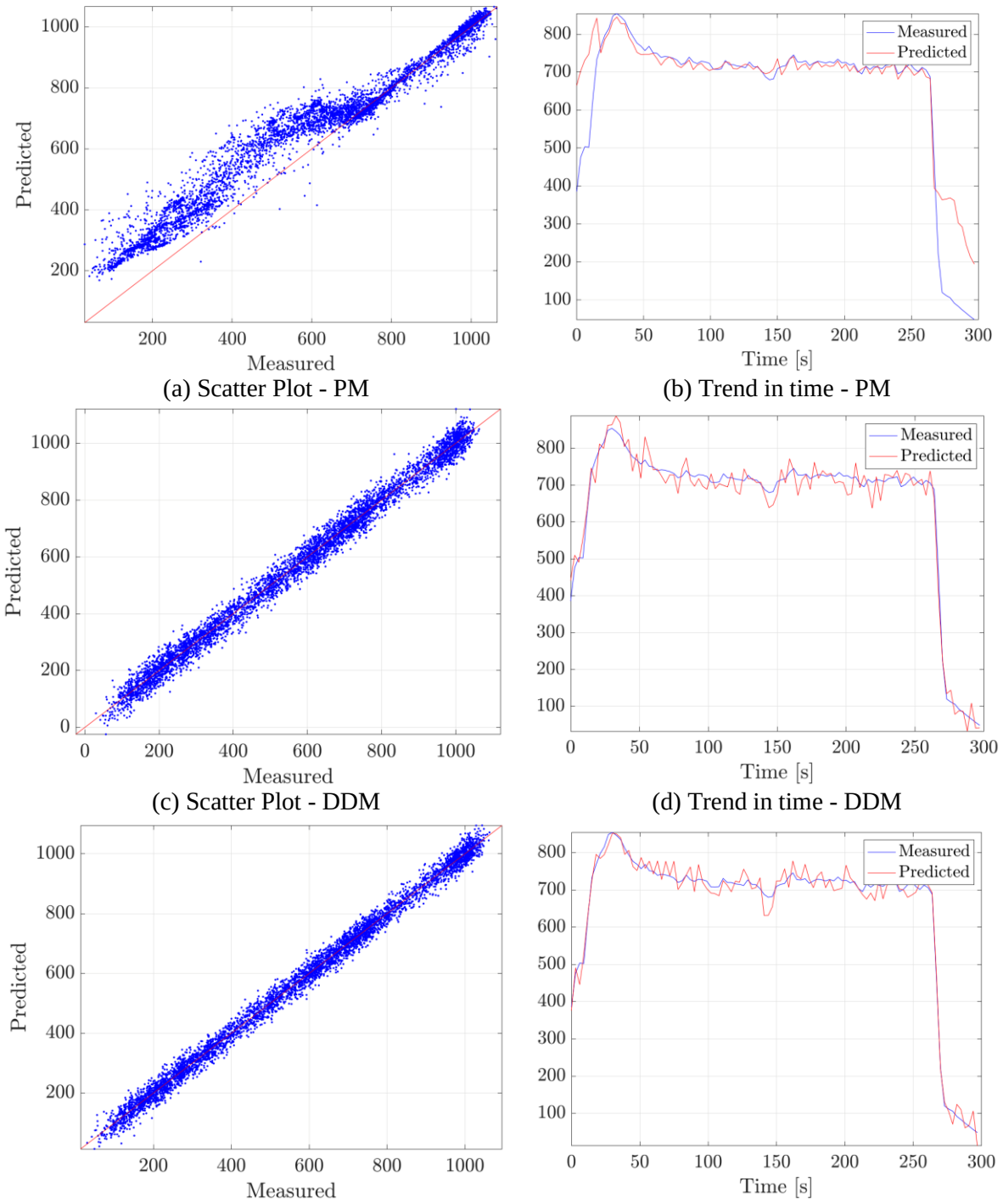
Table III: Table captions above table

Model	MAE [°C]	MAPE [%]	PPMCC	MAE [°C]	MAPE [%]	PPMCC
	Fuel Consumption $\dot{m}_f$ [kg/h]			Turbine Outlet Temperature $T_{\mathrm{t,out}}$ [°C]		
PM	76.62 ± 4.37	26.93 ± 1.54	0.98 ± 0.01	9.66 ± 0.57	2.53 ± 0.13	0.92 ± 0.01
DDM	24.11 ± 1.39	6.30 ± 0.38	0.99 ± 0.01	3.80 ± 0.20	0.97 ± 0.05	0.99 ± 0.01
HM	18.64 ± 0.98	4.89 ± 0.17	1.00 ± 0.01	3.18 ± 0.22	0.81 ± 0.05	0.99 ± 0.01
	TC Rotational speed N_{tc} [rpm]			Exhaust Manifold Temperature T_{er} (°C)		
PM	2090 ± 78.43	15.39 ± 0.75	0.97 ± 0.01	19.92 ±1.06	4.81 ± 0.15	0.96 ± 0.01
DDM	302.6 ± 21.42	2.18 ± 0.15	1.00 ± 0.01	5.02 ± 0.19	1.13 ± 0.04	0.99 ± 0.01
HM	214.44 ± 9.54	1.53 ± 0.08	1.00 ± 0.01	3.94 ± 0.24	0.88 ± 0.05	0.99 ± 0.01

A substantial decrease in the errors can be observed from Table III across all the targets. Considering $\dot{m}_f$, we can observe a MAPE decrease from 26.93% (PM) to 6.30% (DDM), to 4.89% (HM). The same general trend can be reported for $N_{\mathrm{tc}}, T_{\mathrm{t,out}}$, and T_{er}.

Figs.2-5 report the scatter plot and examples of the trend in time for the different targets using PM, DDMs, and HMs. Compared to the PM, the proposed DDMs are more accurate in predicting the four targets. In addition, it is possible to observe that DDMs are capable of fully capturing the transient behaviour of the fuel consumption, Fig.2d, the turbocharger rotational speed mechanical transient, Fig.3d, and the thermodynamic transients of both the turbine outlet gases, Fig.4d, and exhaust manifold, Fig.5d. Also, Figs.2-5 show that the DDMs are characterised by both lower bias and lower variance compared to the PM.

Although DDMs are computationally demanding in the training phase, they are characterised by lower computational complexity in the feed-forward phase as they just require matrix manipulation methods. The combination of both accurate and fast predictions makes DDMs an ideal candidate for real-time performance and condition estimation. However, the necessary data to reach this level of performance is rather high, as reported in *Cipollini et al. (2018a, 2018b)*, making this type of model applicable only after extensive measurement campaigns have been undertaken. In addition, another disadvantage of DDMs is the lack of interpretability as it is not supported by any physical interpretation. To overcome those limitations, we proposed the use of HMs. These allow the exploitation of both the mechanistic knowledge of the underlying physical principles from the PM and any available measurements taken during the operation of the vessel. An advantage of the HMs is their ability to exploit the coarse but physically supported predictions of the PM. Therefore, HMs have much smaller requirements regarding the use of actual measurements for the learning phase. While they will still require a measurement campaign to be deployed, they can be reliably used already after a few months worth of measurements, in contrast with pure DDMs that would require at least half a year of available data.



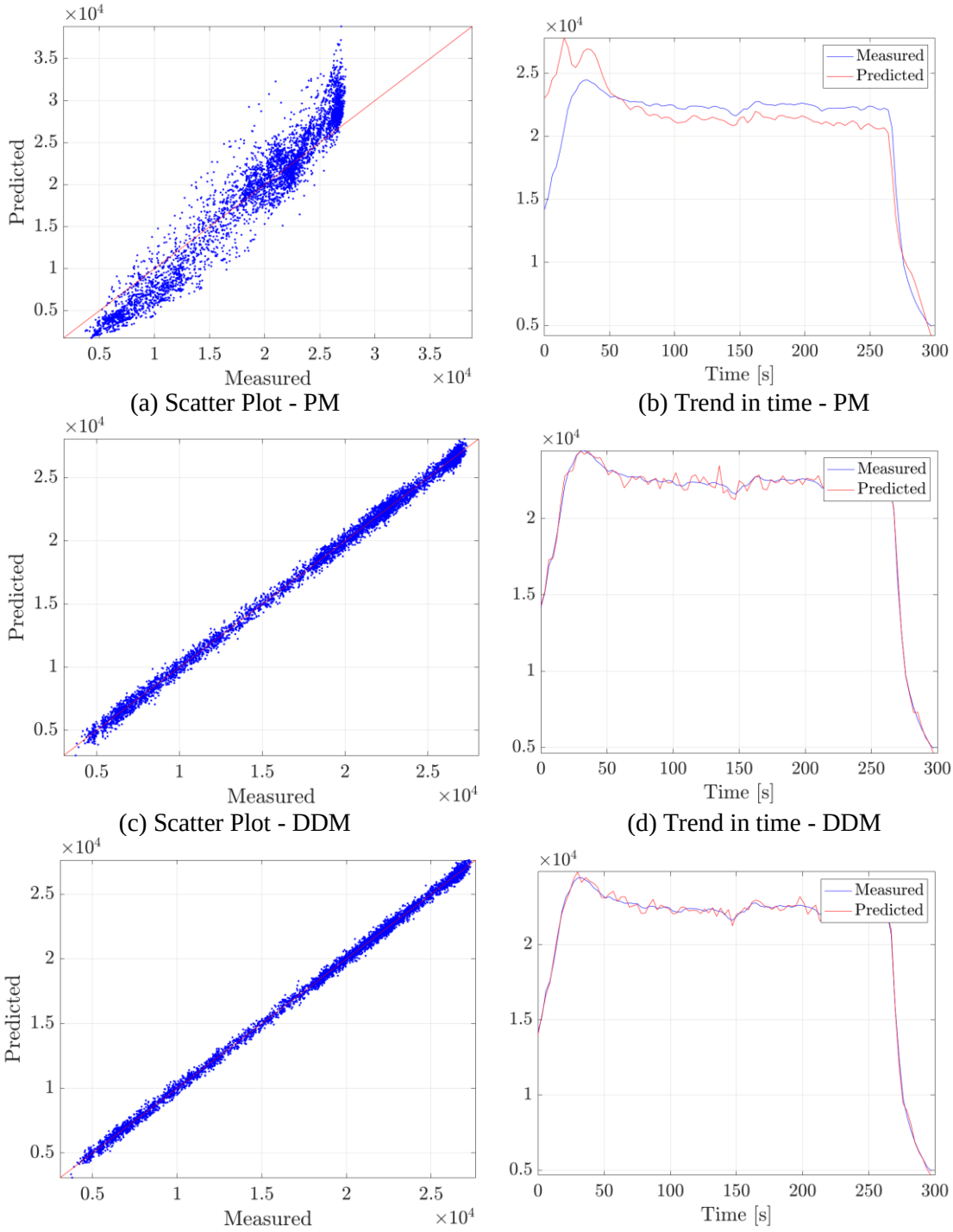
(a) Scatter Plot - PM (b) Trend in time - PM

(c) Scatter Plot - DDM (d) Trend in time - DDM

(e) Scatter Plot - HM (f) Trend in time - HM

Fig.2: Scatter plot and trend in time for the $\dot{m}_f$ (kg/h) output feature - PMs, DDMs, and HMs.

The novelty introduced by the HMs led to more accurate predictions of the four targets compared to the rest of the models (PM and DDMs), as can be seen from Table III.

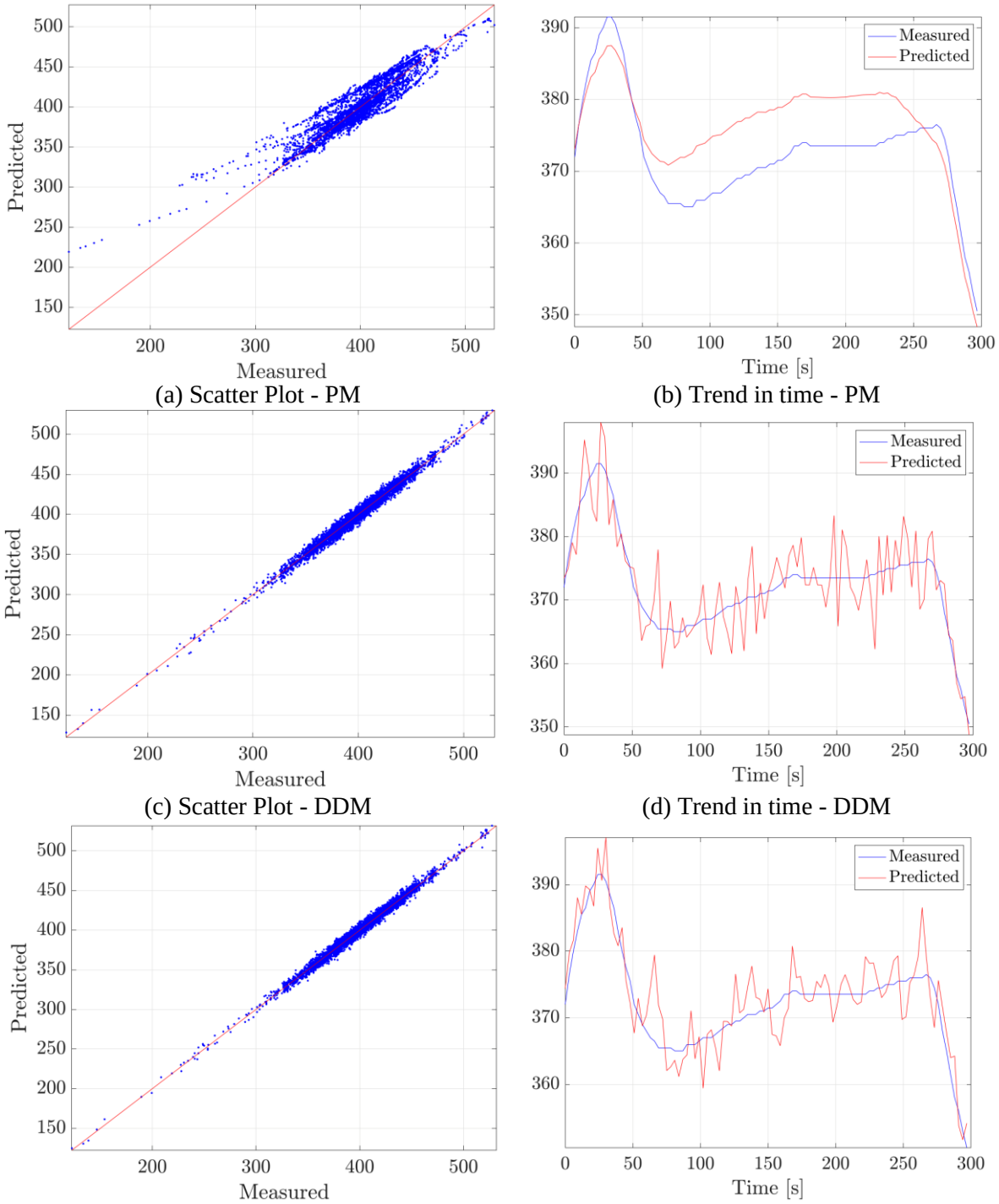


(a) Scatter Plot - PM (b) Trend in time - PM

(c) Scatter Plot - DDM (d) Trend in time - DDM

(e) Scatter Plot - HM (f) Trend in time - HM

Fig.3: Scatter plot and trend in time for the N_{tc} (rpm) output feature - PMs, DDMs, and HMs.

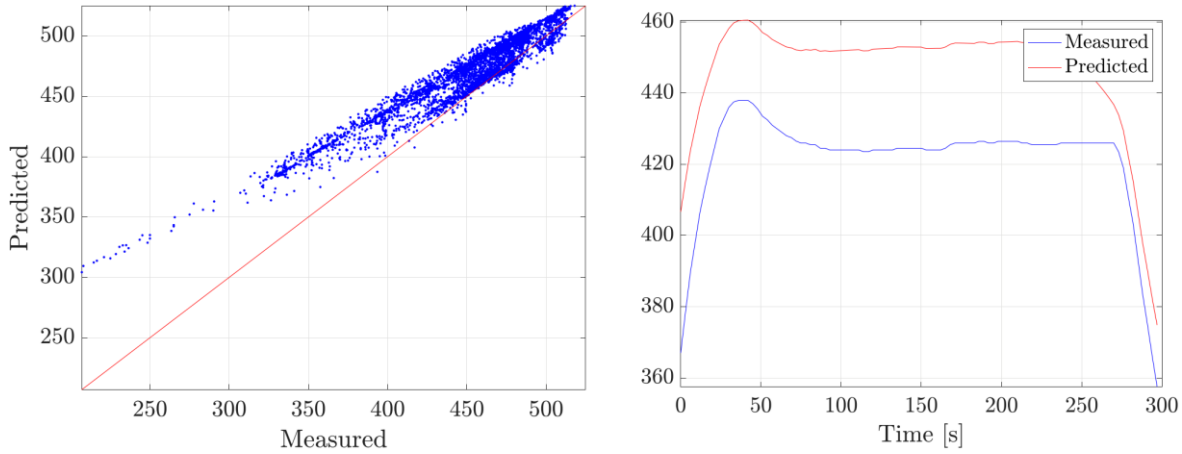


(a) Scatter Plot - PM (b) Trend in time - PM

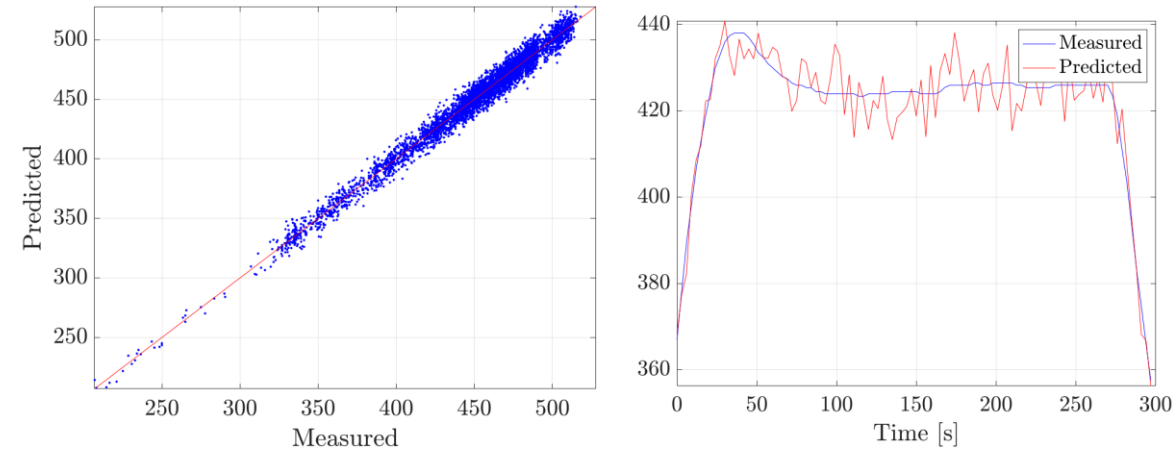
(c) Scatter Plot - DDM (d) Trend in time - DDM

(e) Scatter Plot - HM (f) Trend in time - HM

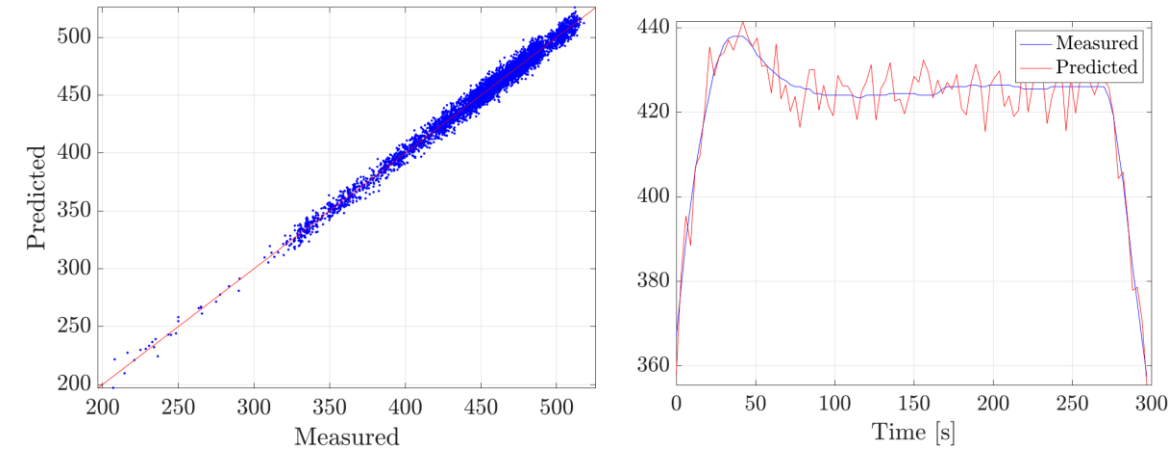
Fig.4: Scatter plot and trend in time for the $T_{t,out}$ (°C) output feature - PMs, DDMs, and HMs.



(a) Scatter Plot - PM (b) Trend in time - PM



(c) Scatter Plot - DDM (d) Trend in time - DDM



(e) Scatter Plot - HM (f) Trend in time - HM

Fig.5: Scatter plot and trend in time for the T_{er} (°C) output feature - PMs, DDMs, and HMs.

6. Conclusions

In this work, the authors focused their attention on demonstrating a novel modelling framework for the hybridisation of physical and data driven models. The proposed framework can deliver accurate, reliable, and computationally inexpensive models suitable for real-time performance assessment and condition monitoring applications. State-of-the-art data-driven methods have been presented, able to exploit the information provided by on-board measurements from one Holland Class Oceangoing Patrol Vessel, provided by the Royal Netherlands Navy and Damen Schelde Naval Shipbuilding. First, a 0D physical model of a medium speed two-stroke diesel

engine (MAN 12 V28-33D) was described. Data-driven models have been discussed and proposed in Section 3.1 to predict the engine's behaviour, with a focus on four different targets: fuel consumption, turbocharger rotational speed, turbine outlet temperature, and exhaust receiver temperature. The models proved to be very accurate, with the enhanced capability of exploiting time-series information from the past, achieving relative errors below 1% on the validation data for the turbine outlet temperature and exhaust receiver temperature. However, due to their nature, these data-driven models are hard to interpret. To overcome the limitations of both the physical and the data-driven models, we proposed a hybrid approach that can consider past information, capable of improving accuracy, easily interpreted, and have low computational time requirements. The hybridisation of physical and data driven models proved to be highly accurate, achieving even lower errors when compared to the simple data-driven approach. These hybrid models can potentially also be used to improve the accuracy of predictions for operation in other conditions than the measured ones, as purely data-driven models cannot be used for extrapolation, but the physical model contribution will improve hybrid model performance during extrapolation.

References

ARGYRIOU, A.; EVGENIOU, T.; PONTIL, M. (2008), *Convex multi-task feature learning,* Machine Learning 73, pp.243-272

ASAD, U.; TJONG, J.; ZHENG, M. (2014), *Exhaust gas recirculation-zero dimensional modelling and characterisation for transient diesel combustion control,* Energy Conversion and Management 86, pp.309-324

BAKKER, B.; HESKES, T. (2003), *Task clustering and gating for Bayesian multitask learning,* J. Machine Learning Research 4, pp.83-99

BALDI, F.; JOHNSON, H.; GABRIELII, C.; ANDERSSON, K. (2014), *Energy and exergy analysis of ship energy systems-the case study of a chemical tanker,* Int. Conf. on Efficiency, Cost, Optimisation, Simulation and Environmental Impact of Energy Systems

BALDI, F.; THEOTOKATOS, G.; ANDERSSON, K. (2015), *Development of a combined mean value-zero-dimensional model and application for a large marine four-stroke diesel engine simulation.* Applied Energy 154, pp.402-415

BAXTER, J. (2000), *A model of inductive bias learning,* J. Artificial Intelligence Research 12, pp.149–198

BIDARVATAN, M.; THAKKAR, V.; SHAHBAKHTI, M.; BAHRI, B.; AZIZ, A. (2014), *Grey-box modeling of HCCI engines,* Applied Thermal Engineering 70, pp.397-409

BONDARENKO, O.; FUKUDA, T. (2020), *Development of a diesel engine's digital twin for predicting propulsion system dynamics,* Energy 196, 117126

CARUANA, R. (1997), *Multitask learning,* Machine Learning 28, pp.41-75.

CASEY, M.; ROBINSON, C. (2013), *A method to estimate the performance map of a centrifugal compressor stage*, J. Turbomachinery 135

CATANIA, A.; FINESSO, R.; SPESSA, E. (2011), *Predictive zero-dimensional combustion model for di diesel engine feed-forward control,* Energy Conversion and Management 52, pp.3159-3175

CIPOLLINI, F.; ONETO, L.; CORADDU, A.; MURPHY, A.; ANGUITA, D. (2018a), *Condition-based maintenance of naval propulsion systems: Data analysis with minimal feedback,* Reliability Engineering & System Safety 177, pp.12-23

CIPOLLINI, F.; ONETO, L.; CORADDU, A.; MURPHY, A.; ANGUITA, D. (2018b), *Condition-based maintenance of naval propulsion systems with supervised data analysis,* Ocean Eng. 149, pp.268–278

EUROPEAN COMMISSION (2013a), *Integrating maritime transport in the EU's green house gas reduction policies: Communication from the commission to the European parliament, the council, the European economic and social committee and the committee of the regions,* Tech. rep. European Union

IMO (2011), *Amendments to the annex of the protocol of 1997 to amend the international convention for the prevention of pollution from ships, 1973, as modified by the protocol of 1978 relating thereto,* Resolution MEPC. 203 (62)

CORADDU, A.; KALIKATZARAKIS, M.; ONETO, L.; MEIJN, G.J.; GODJEVAC, M.; GEERTS-MAD, R.D. (2018), *Ship diesel engine performance modelling with combined physical and machine learning approach,* Int. Naval Engineering Conf.

CORADDU, A.; LIM, S.; ONETO, L.; PAZOUKI, K.; NORMAN, R.; MURPHY, A. (2019a), *A novelty detection approach to diagnosing hull and propeller fouling,* Ocean Eng. 176, pp.65–73

CORADDU, A.; ONETO, L.; BALDI, F.; ANGUITA, D. (2017), *Vessels fuel consumption forecast and trim optimisation: a data analytics perspective,* Ocean Eng. 130, pp.351–370

CORADDU, A.; ONETO, L.; BALDI, F.; CIPOLLINI, F.; ATLAR, M.; SAVIO, S. (2019b), *Data-driven ship digital twin for estimating the speed loss caused by the marine fouling*, Ocean Eng. 186, 106063

CORADDU, A.; ONETO, L.; CIPOLLINI, F.; KALIKATZARAKIS, M.; MEIJN, G. J.; GEERTSMA, R. (2021a), *Physical, data-driven and hybrid approaches to model engine exhaust gas temperatures in operational conditions,* Ships and Offshore Structures, pp.1–22

CORADDU, A.; ONETO, L.; ILARDI, D.; STOUMPOS, S.; THEOTOKATOS, G. (2021b), *Marine dual fuel engines monitoring in the wild through weakly supervised data analytics,* Engineering Applications of Artificial Intelligence 100, 104179

CORADDU, A.; ONETO, L.; DE MAYA, B.; KURT, R. (2020*), Determining the most influential human factors in maritime accidents: A data-driven approach,* Ocean Eng. 211, 107588

CRISTIANINI, N.; SHAWE-TAYLOR, J. (2000), *An introduction to support vector machines and other kernel-based learning methods,* Cambridge University Press

DING, Y.; STAPERSMA, D.; KNOLL, H.; GRIMMELIUS, H.; NETHERLAND, T. (2010). *Characterising heat release in a diesel engine: A comparison between seiliger process and vibe model,* CIMAC International Council on Combustion engines, pp.1–13

EVGENIOU, T.; PONTIL, M. (2004), *Regularised multi-task learning,* Int. Conf. on Knowledge Discovery and Data Mining

FERNÁNDEZ-DELGADO, M.; CERNADAS, E.; BARRO, S.; AMORIM, D. (2014), *Do we need hundreds of classifiers to solve real world classification problems?,* J. Machine Learning Research 15, pp.3133–3181

GARCÍA-MARTOS, C.; RODRÍGUEZ, J.; SÁNCHEZ, M. J. (2013), *Modelling and forecasting fossil fuels, CO2 and electricity prices and their volatilities,* Applied Energy 101, pp.363–375

GEERTSMA, R.; NEGENBORN, R.; VISSER, K.; LOONSTIJN, M.; HOPMAN, J. (2017), *Pitch control for ships with diesel mechanical and hybrid propulsion: Modelling, validation and performance quantification,* Applied Energy 206, pp.1609–1631

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. (2016), *Deep Learning*, MIT Press

GRIMMELIUS, H.; BOONEN, E.; NICOLAI, H.; STAPERSMA, D. (2010), *The integration of mean value first principle diesel engine models in dynamic waste heat and cooling load analysis,* CIMAC Congress, Bergen, Vol 31, p.51

GUAN, C.; THEOTOKATOS, G.; ZHOU, P.; CHEN, H. (2014), *Computational investigation of a large containership propulsion engine operation at slow steaming conditions,* Applied Energy 130, pp.370–383

HAMILTON, J.D. (2020), *Time series analysis*, Princeton University Press

HANSON, R.K.; SALIMIAN, S. (1984), *Survey of rate constants in the N/H/O system*, Combustion Chemistry, pp.361–421

HEYWOOD, J.B. (1988), *Internal combustion engines fundamentals*, McGraw-Hill

KALIKATZARAKIS, M.; CORADDU, A.; THEOTOKATOS, G.; ONETO, L. (2021), *Development of a zero-dimensional model and application on a medium-speed marine four-stoke diesel engine,* Int. Conf. on Modelling and Optimisation of Ship Energy Systems

KEERTHI, S.S.; LIN, C.J. (2003), *Asymptotic behaviors of support vector machines with gaussian kernel*, Neural computation 15, pp.1667–1689

LIVANOS, G.; PAPALAMBROU, G.; KYRTATOS, N.; CHRISTOU, A. (2007), *Electronic engine control for ice operation of tankers*, 25th CIMAC World Congress on Combustion Engine Technology, Vienna, pp.21–24

MERKER, G.P.; SCHWARZ, C.; STIESCH, G.; OTTO, F. (2005), *Simulating Combustion: Simulation of combustion and pollutant formation for engine developmen*t, Springer Science & Business Media

MIGLIANTI, F.; CIPOLLINI, F.; ONETO, L.; TANI, G.; VIVIANI, M. (2019), *Model scale cavitation noise spectra prediction: Combining physical knowledge with data science*, Ocean Eng. 178, pp.185–203

MIGLIANTI, L.; CIPOLLINI, F.; ONETO, L.; TANI, G.; GAGGERO, S.; CORADDU, A.; VIVIANI, M. (2020), *Predicting the cavitating marine propeller noise at design stage: A deep learning based approach*, Ocean Eng. 209, 107481

MISHRA, C.; SUBBARAO, P. (2021), *A Comparative Study of Physics Based Grey Box and Neural Network Trained Black Box Dynamic Models in an RCCI Engine Control Parameter Prediction,* Technical Report SAE Technical Paper

MOHAMMADKHANI, F.; YARI, M.; RANJBAR, F. (2019), *A zero-dimensional model for simulation of a Diesel engine and exergoeconomic analysis of waste heat recovery from its exhaust and coolant employing a high-temperature Kalina cycle*, Energy Conversion and Management 198, 111782

NIKZADFAR, K.; SHAMEKHI, A. H. (2014), *Investigating the relative contribution of operational*

parameters on performance and emissions of a common-rail diesel engine using neural network, Fuel 125, pp.116–128

ONETO, L. (2020), *Model Selection and Error Estimation in a Nutshell*, Springer

ONETO, L.; GHIO, A.; RIDELLA, S.; ANGUITA, D. (2015), *Support vector machines and strictly positive definite kernel: The regularisation hyper parameter is more important than the kernel hyperparameters*, IEEE Int. Joint Conf. on Neural Networks

ÖZENER, O.; YUKSEK, L.; OZKAN, M. (2013), *Artificial neural network approach to predicting engine-out emissions and performance parameters of a turbo charged diesel engine*, Thermal Science 17, pp.153–166

RAKOPOULOS, C.D.; HOUNTALAS, D.T.; TZANOS, E.I.; TAKLIS, G.N. (1994), *A fast algorithm for calculating the composition of diesel combustion products using 11 species chemical equilibrium scheme*, Advances in Engineering Software 19, pp.109–119

ROHSENOW, W.M.; HARTNETT, J.P. (1988), *Handbook of heat transfer fundamentals*, McGraw-Hill

ROSASCO, L.; DE VITO, E.; CAPONNETTO, A.; PIANA, M.; VERRI, A. (2004), *Are loss functions all the same*?, Neural Computation 16, pp.1063–1076

SAPRA, H.; GODJEVAC, M.; DE VOS, P.; VAN SLUIJS, W.; LINDEN, Y.; VISSER, K. (2020), *Hydrogen-natural gas combustion in a marine lean-burn SI engine: A comparitive analysis of Seiliger and double Wiebe function-based zero–dimensional modelling*, Energy Conversion and Management 207, 112494

SCHÖLKOPF, B. (2001), *The kernel trick for distances*, Advances in Neural Information Processing Systems, pp.301–307

SCHÖLKOPF, B.; HERBRICH, R.; SMOLA, A.J. (2001), *A generalized representer theorem*, Computational Learning Theory

SHALEV-SHWARTZ, S.; BEN-DAVID, S. (2014), *Understanding machine learning: From theory to algorithms*, Cambridge University Press

SHAO, L.; MAHAJAN, A.; SCHRECK, T.; LEHMANN, D.J. (2017), *Interactive regression lens for exploring scatter plots*, Computer Graphics Forum

SHAWE-TAYLOR, J.; CRISTIANINI, N. (2004), *Kernel methods for pattern analysis*, Cambridge University Press

SYED, J.; BAIG, R. U.; ALGARNI, S.; MURTHY, S.; MASOOD, M.; INAMURRAHMAN, M. (2017), *Artificial Neural Network modeling of a hydrogen dual fueled diesel engine characteristics: An experiment approach*, Int. J. Hydrogen Energy 42(21), pp.14750-14774

TIKHONOV, A.N.; ARSENIN, V.Y. (1979), *Methods for the Solution of Ill-Posed Problems*, Nauka

VAPNIK, V.N. (1998), *Statistical learning theory*, Wiley

VOVK, V. (2013), *Kernel ridge regression*, Empirical Inference

WAINBERG, M.; ALIPANAHI, B.; FREY, B.J. (2016), *Are random forests truly the best classifiers?*, J. Machine Learning Research 17, pp.3837–3841

WATSON, N.; JANOTA, M. (1982), *Turbocharging the internal combustion engine*, Macmillan Int. Higher Education

WILLMOTT, C.J.; MATSUURA, K. (2005), *Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance*, Climate Research 30, pp.79-82

WOSCHNI, G. (1967), *A universally applicable equation for the instantaneous heat transfer coefficient in the internal combustion engine*, Technical Report SAE Technical paper

YOUNG, D.M. (2003), *Iterative solution of large linear systems*, Dover Publications