

Ship diesel engine performance modelling with combined physical and machine learning approach

A. Coraddu, PhD, NAOME^{a*}, M. Kalikatzarakis, MSc^b, L. Oneto, PhD, DIBRIS^c, MSc, G.J. Meijn, MSc^b, M. Godjevac, PhD^d, R.D. Geertsma^{d,e*}

^aDepartment Of Naval Architecture, Ocean & Marine Engineering - University of Strathclyde - UK;

^bResearch & Technology Support - Damen Schelde Naval Shipbuilding - the Netherlands

^cDIBRIS - University of Genoa - Italy

^dDepartment of Maritime & Transport Technology - Delft University of Technology - the Netherlands

^eFaculty of Military Sciences - Netherlands Defence Academy - the Netherlands

*Corresponding author. Email: r.d.geertsma@tudelft.nl

Synopsis

Condition Based Maintenance on diesel engines can help to reduce maintenance load and better plan maintenance activities in order to support ships with reduced or no crew. Diesel engine performance models are required to predict engine performance parameters in order to identify emerging failures early on and to establish trends in performance reduction. In this paper, a novel approach is proposed to accurately predict engine temperatures during operational dynamic manoeuvring. In this hybrid modelling approach, the authors combine the mechanistic knowledge from physical diesel engine models with the statistic knowledge from engine measurements on a sound engine. This simulation study, using data collected from a *Holland* class patrol vessel, demonstrates that existing models cannot accurately predict measured temperatures during dynamic manoeuvring, and that the hybrid modelling approach outperforms a purely data driven approach by reducing the prediction error during a typical day of operation from 10% to 2%.

Keywords: Data-Drive methods; Condition Based Maintenance; Gray-Box Models; Exhaust gas temperature prediction; Machine Learning

1 Introduction

The Royal Netherlands Navy (RNLN) in 2012 introduced the Oceangoing Patrol Vessels (OPVs), shown in Figure 1, to perform worldwide operations at the lower end of the warfare spectrum, such as counter drugs, counter piracy, coastguard and disaster relief operations, with a crew of only 50. This could be achieved due to a high degree of innovation (Horenberg and Melaet, 2013), advanced firefighting and damage control automation (Geertsma et al., 2013a) and Combat and Platform Management System (PMS) integration (Geertsma et al., 2013b). However, Horenberg and Melaet (2013) demonstrate that the maintenance load for the OPV is too high for its crew and shore support organisation and propose Condition Based Maintenance (CBM) as a measure to reduce maintenance load.

The increased level of automation has led to an increase in available sensor information to be used for CBM, particularly for diesel engines. Moreover, diesel engines require a significant amount of maintenance on the OPV class (Horenberg and Melaet, 2013). Recent advances in *Machine Learning* techniques allow the use of sensor data to indicate the degradation in performance of naval gas turbine engines (Coraddu et al., 2016). Moreover, Coraddu et al. (2017) have demonstrated the potential of using a combination of *Machine Learning* algorithms and physical models in predicting vessel fuel consumption and determining optimal trim. However, this combined approach has not yet been applied to improve diesel engine models, so they can be used for CBM.

Authors' Biographies

Dr Andrea Coraddu currently is Lecturer at the Department of Naval Architecture, Ocean & Marine Engineering at the University of Strathclyde. His research interests focus on improving the understanding, design, and performance of propulsion plants simulations and control, mainly through the application of numerical models, data mining, statistics, and performance evaluation.

Prof. Luca Oneto was born in Rapallo, Italy in 1986. He received his BSc and MSc in Electronic Engineering at the University of Genoa, Italy respectively in 2008 and 2010. In 2014 he received his PhD from the same university in School of Sciences and Technologies for Knowledge and Information Retrieval with the thesis "Learning Based On Empirical Data". In 2017 he obtained the Italian National Scientific Qualification for the role of Associate Professor in Computer Engineering and in 2018 the one in Computer Science. He is currently an Assistant Professor in Computer Engineering at University of Genoa with particular interests in Statistical Learning Theory and Data Science.

Lt Cdr (E) Rinze Geertsma, MSc, CEng, currently is PhD researcher at Delft University of Technology and Netherlands Defence Academy on control strategies for hybrid ship systems. He has previously been Marine Engineering Officer of *HNLMS de Ruyter* and *HNLMS Tromp*. Earlier experience includes research and development project managing, system and project engineering and in service support.

Gert-Jan Meijn, MSc, currently is research engineer at Damen Schelde Naval Shipbuilding. He received his BSc. and MSc. in Mechanical Engineering at Delft University of Technology in 2013 and 2015 respectively. His working experience includes dynamic simulations of propulsion systems and performing propulsion configuration trade-off studies.

Miltos Kalikatzarakis, MSc, currently is research engineer at Damen Schelde Naval Shipbuilding. He received his MEng. in Naval Architecture at the National Technical University of Athens, and MSc. in Marine Technology (Marine engineering) at Delft University of Technology in 2013 and 2017 respectively.

Dr Milinko Godjevac received his PhD degree in marine engineering from Delft University of Technology in 2010 and currently works as assistant professor at the Department of Maritime & Transport Technology in Delft University of Technology. His research is focused on integration and control of on-board power systems.



Figure 1: RNLN Holland class Oceangoing Patrol Vessel, case study in this paper.

Currently, most diesel engine models focus on the prediction of fuel consumption, CO₂, NO_x and PM emissions and smoke production (Theotokatos and Tzelepis, 2015; Nielsen et al., 2017; Kökkülünk et al., 2016). However, temperatures of the exhaust valve, exhaust receiver gas flow and turbine entry gas flow are important indicators for thermal loading of an engine (Sapra et al., 2017; Geertsma et al., 2017; Grimmeliuss and Stapersma, 2000), and to indicate intercooler failure (Hountalas, 2000). The current state of the art models still have a large error of up to 6% in predicting temperatures in static conditions (Geertsma et al., 2017; Theotokatos and Tzelepis, 2015), while dynamic predictions are worse, as shown in this paper. The necessity for accurate dynamic temperature predictions, somehow bypassing the thermal inertia of the sensors themselves, is demonstrated by Korczewski (2016). Deviation of measured temperatures from temperatures predicted by a model that takes engine operating conditions into account can provide important diagnostic information to detect deterioration of exhaust valves and turbocharger stator and rotor blades before failure (Korczewski, 2015).

With the rapid growth of ship monitoring systems within the shipping industry, the use of *Data-Driven Models* is obtaining substantial interest. Models based on the mechanical understanding of the physics of the problem are well-established, although they could fail in the accurate forecast of the effect of various distinct and heterogeneous phenomena. On the other hand, *Data-Driven Models* can be applied by employing a large amount of data with few, if no physical knowledge of the system. Lately, *Machine Learning* techniques have been successfully employed for the evaluation of ship performance (Coraddu et al., 2015; Petersen et al., 2012). Artificial Neural Network and Gaussian process methods were applied to estimate the ship's fuel efficiency in Pedersen and Larsen (2009) and Petersen et al. (2012), and in Radonjic and Vukadinovic (2015) an Ensemble Neural Network was applied to predict a towboat shaft power demand. In Petersen et al. (2012) a statistical model for the fuel consumption prediction, adopting Artificial Neural Networks and Gaussian Processes is proposed, and finally, in Coraddu et al. (2017) the authors showed how *Data-Driven Models* outperform physical models in the prediction of ship performance. In this paper, an existing diesel engine model is improved using data-driven techniques adding a black-box component to the model output (Leifsson et al., 2008; Coraddu et al., 2017). The proposed gray-box modelling approach allows exploiting both the mechanistic knowledge of the physical principles and available measurements. As reported in Coraddu et al. (2017) this approach provides more accurate outcomes when compared with the first principle physical model, and requires a smaller amount of data when compared to the *Data-Driven Models*. Therefore, this work aims to investigate how *Data-Driven Models* and a combination of data-driven and physical models can improve prediction of engine temperatures, using extensive PMS measurement data from the OPV class.

The novelties of this paper are: first, the authors use data analysis on an extensive dataset to establish whether *Data-Driven Models* can be used to predict engine temperatures; secondly, the authors present a novel gray-box modelling approach to predict diesel temperatures using the physical models from Geertsma et al. (2017) in combination with *Machine Learning* algorithms from Coraddu et al. (2016, 2017). The resulting *Gray-Box Model* should yield a more accurate representation of the engine, which will then be suitable for use in various aspects of off-line and real-time operational optimisation.

The paper is organised as follows: Section 2 gives a brief description of the system, the used physical model and the dataset used for this work, Section 4 describes the *Data-Driven Model* and associated theory, Section 5 describes the combined physical and data learning *gray-box* approach. Finally, Section 6 describes the results of the three modelling approaches and Section 7 summarises the conclusions and proposes future work.

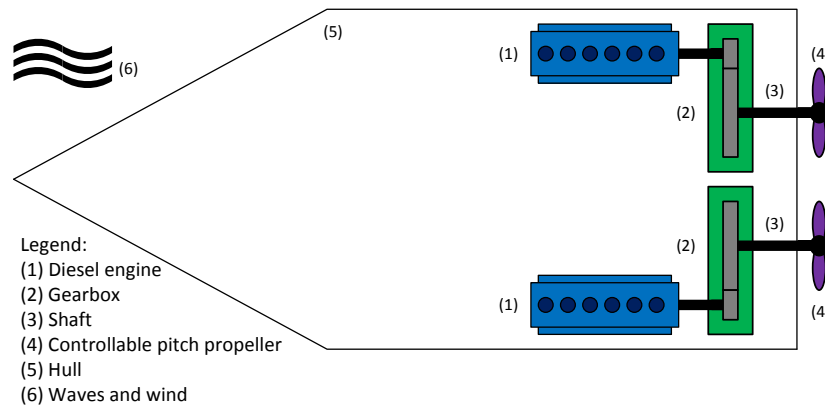


Figure 2: Propulsion system layout for the considered naval vessel, from Geertsma et al. (2017).

Variable name	Unit	Variable name	Unit
Drive shaft torque	[Nm]	HT cooling water temperature after cooler	[°C]
Engine speed	[rpm]	HT cooling water temperature before cooler	[°C]
Main bearing #1 temperature	[°C]	Relative charge air pressure	[bar]
Main bearing #2 temperature	[°C]	Charge air temperature after Turbocharger A bank	[°C]
Main bearing #3 temperature	[°C]	Charge air temperature after Turbocharger B bank	[°C]
Main bearing #4 temperature	[°C]	Charge air temperature before Turbocharger A bank	[°C]
Main bearing #5 temperature	[°C]	Charge air temperature before Turbocharger B bank	[°C]
Main bearing #6 temperature	[°C]	Temperature charge air	[°C]
Main bearing #7 temperature	[°C]	Turbocharger speed TC A bank	[rpm]
Splash oil compartment #1 temperature	[°C]	Turbocharger speed TC B bank	[rpm]
Splash oil compartment #2 temperature	[°C]	Exhaust temperature after TC A bank	[°C]
Splash oil compartment #3 temperature	[°C]	Exhaust temperature after TC B bank	[°C]
Splash oil compartment #4 temperature	[°C]	Fuel temperature	[°C]
Splash oil compartment #5 temperature	[°C]	LT cooling water temperature after cooler	[°C]
Splash oil compartment #6 temperature	[°C]	LT cooling water temperature before cooler	[°C]
Fuel rack position	[%]	Power margin relative to maximum	[%]
Lube oil temperature before engine	[°C]	Average exhaust gas temperature A bank	[°C]
Lube oil temperature after engine	[°C]	Average exhaust gas temperature b bank	[°C]
Fuel flow	[kg/s]		

Table 1: System’s main features.

2 System physical model description and available data

The propulsion system of the Holland class Patrol Vessels consists of two shafts with controllable pitch propellers (CPP), a gearbox, and one diesel engine per shaft, as shown in Figure 2. In this study, the authors consider the performance of the engine without taking into account the interaction with gearbox, propeller and ship, because this interaction influences engine behaviour primarily through engine load, a measured output.

For the *Engine Physical Model*, *Engine Data-Driven Model* and the combined *Engine Hybrid Model*, the authors use the dataset of one four-stroke, medium speed diesel engine on board a *Holland class* Oceangoing Patrol Vessel, which is shown in Figure 1. The vessel is equipped with a data logging system which is used by the Royal Netherlands Navy both for on board monitoring and for land-based performance control. The dataset consists of 114 signals with a sample rate of 3 s that cover a time of 35 hours, totalling 42,000 data points. In brief, these points consist of several control and monitoring parameters of the engine, from engine speed and torque, to various operational pressures and temperatures of engine components such as the crankshaft, cylinder and turbocharger and systems, such as water cooling, lubricating oil, exhaust-gas, and fuel systems. Table 1 summarises the subset of the available measurements, from the continuous monitoring system, that have been used in the *Machine Learning* modelling phase.

3 Engine Physical Model

The *Engine Physical Model* is a mean value first principle diesel engine model, illustrated in Figure 3, and described in detail in Geertsma et al. (2017). The model consists of three state variables: fuel injection per cylinder per cycle m_f , charge pressure p_1 and exhaust receiver pressure p_d . The inputs of the model are engine speed n_e and fuel pump setpoint X_{set} , which typically originates from the speed governor, and the output is engine torque M_e .

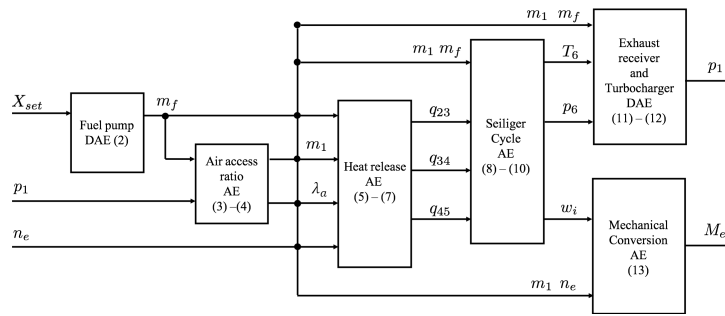


Figure 3: Schematic representation of the diesel engine model and the interaction of its subsystems, from Geertsma et al. (2017).

The temperatures of the gas flow in the exhaust receiver and at the turbocharger exit, subjects of this study, are represented by a system of Algebraic Equations (AE) and Differential and Algebraic Equations (DAE) featuring the input variables, state variables and the following mathematically related parameters: trapped mass in the cylinder m_1 , air excess ratio λ_a , isobaric, isochoric and isothermal heat release q_{23} , q_{34} and q_{45} , temperature and pressure after expansion of the Seiliger cycle T_6 and p_6 and induced work during the Seiliger cycle w_i . The original aspect of this model is that the turbocharger dynamics are represented by the Büchi power and flow balance between compressor and turbine and do not require compressor or turbine maps for calibration. As fast dynamics are neglected in this model, the model runs up to 2500 times real time, much faster than mean value engine models using compressor and turbine maps, such as Nielsen et al. (2017), Theotokatos and Tzelepis (2015), Sapra et al. (2017) and Kökkülünk et al. (2016). Moreover, *Machine Learning* methods can be used to improve the mathematical relationships in the models, improving its accuracy, while maintaining the run-time during the feed-forward phase, after an initial learning phase, which also benefits from fast physical models.

The *Engine Physical Model* predicts gas flow temperatures, in particular the temperatures after the turbine exit T_e and the exhaust receiver temperature T_d , and neglects the effect of the location of the thermocouples, which measure the temperature at the outside of the piping. Therefore, to account for the location of the sensors, the authors have included thermal inertia and heat losses with a first order response with a constant time constant of the thermocouple τ_{tc} and a constant bias b_{tc} , as follows:

$$\frac{T_{tc,d}(t)}{dt} = \frac{(T_d(t) - b_{tc,d}) - T_{tc,d}(t)}{\tau_{tc,d}} \tag{1}$$

$$\frac{T_{tc,e}(t)}{dt} = \frac{(T_e(t) - b_{tc,e}) - T_{tc,e}(t)}{\tau_{tc,e}}, \tag{2}$$

in an effort to better align the physical prediction with the actual temperature measurement at turbocharger entry, $T_{tc,d}$ and turbocharger exit $T_{tc,e}$ from the dataset.

4 Engine Data-Driven Model

Machine Learning approaches play a central role in extracting non-trivial information from amounts of raw data, collected from sensors and heterogeneous inputs and concerning the monitored asset. The learning process for *Machine Learning* approaches usually consists of two phases: during the *training phase*, a set of data is used to induce a model that best fits them, according to some criteria; then, the trained model is set in motion into the real application during the *feed-forward phase*.

In this paper, the authors focus on the problem of ship diesel engine performance modelling, developing the so called *Engine Data-Driven Model*. For this model development, a *regression* problem must be solved, where the key temperatures must be estimated based on data, collected from the vessel’s monitoring system. In particular, in the conventional regression framework in *Machine Learning* (Vapnik, 1998; Cristianini and Shawe-Taylor, 2000; Zou and Hastie, 2005), a set of training data $\mathcal{D} = \{(x_1, y_1), \dots, (x_n, y_n)\}$, where $x \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$, is considered. As the authors are targeting a regression problem, the purpose is to find the best approximating function in a set of possible ones $h \in \mathcal{H}$, where $h : \mathbb{R}^d \rightarrow \mathbb{R}$, which should be close (in some sense) to the unknown target function $f(x)$.

During the training phase, the quality of the learned regressor $h(x)$ is usually measured according to a loss function $\ell(h(x), y)$ (Rosasco et al., 2004), which calculates the discrepancy between the true output y and the estimated one. The empirical error then computes the average discrepancy per sample, reported by a model on the

training patterns:

$$\hat{L}(h) = \frac{1}{n} \sum_{i=1}^n \ell(h(x_i), y_i). \tag{3}$$

A simple criterion for selecting the final model during the training phase could then consist in simply choosing the approximating function that minimises the empirical error $\hat{L}(h)$. This approach, known as *Empirical Risk Minimization* (Vapnik, 1998), is usually avoided in *Machine Learning* as it leads to severe overfitting of the model on the training dataset.

A more effective approach is represented by minimising 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 (Tikhonov and Arsenin, 1979):

$$h^* : \min_{h \in \mathcal{H}} \hat{L}(h) + \lambda C(h). \tag{4}$$

In other words, the best approximating function h^* is chosen as the one that is complicated enough to learn from data without overfitting them. In particular, $C(\cdot)$ is a complexity measure: depending on the exploited *Machine Learning* approach, different measures are exploited since no optimal choice exists (Wolpert and Macready, 1997). Instead, λ is a hyperparameter, that must be aprioristically set 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 $C(\cdot)$. The optimal value for λ is problem-dependent, and tuning this hyperparameter is a non-trivial task that must be solved during the *Model Selection* phase (Oneto, 2018). Obviously, the error that h commits over \mathcal{D}_n is optimistically biased since \mathcal{D}_n has been used for building h itself. For this reason, another set of fresh data, composed of m samples and called test set $\mathcal{T}_m = \{(x_1^t, y_1^t), \dots, (x_m^t, y_m^t)\}$, needs to be exploited.

Among the several loss-function available to estimate the performances of the learned regressor $h(x)$ in approximating the target function $f(x)$, the squared loss $\ell_2(f(x), y) = (f(x) - y)^2$ is one of the mostly adopted. The squared loss can be adopted to create different indexes of performance, which differently weight the distance between y_i^t and $f(x_i^t)$, such as the Root Mean Square Error (RMSE) and Relative Error Percentage (REP), where:

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m \ell_2(f(x_i^t), y_i^t)}. \tag{5}$$

$$\text{REP} = 100 \sqrt{\frac{\sum_{i=1}^m \ell_2(f(x_i^t), y_i^t)}{\sum_{i=1}^m (y_i^t)^2}}. \tag{6}$$

In order to solve the problem, the authors have to first define the functional structure of h . It is possible to define h as a nonlinear function (Vapnik, 1998) $h(x) = w \cdot \phi(x) + b$ where $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^D$ with usually $D \gg d$, $w \in \mathbb{R}^D$, and $b \in \mathbb{R}$. Since the authors are dealing with regression problems, the most suited loss function is the square loss function $\ell(h(x), y) = (y - h(x))^2$ (Rosasco et al., 2004). Finally, as a complexity measure, the combination of the L1 and the L2 regularization schemes (Zou and Hastie, 2005) has been exploited. L1L2 regularization schema, also called *Elastic Net* regularization (Zou and Hastie, 2005), is both a regularization and variable selection method. L1L2 often outperforms the L1, while providing a similar sparsity of representation. In addition, the L1L2 encourages a grouping effect, where strongly correlated features tend to be in or out of the model together. Consequently, Problem (4) can be reformulated as follows:

$$(w, b)^* : \arg \min_{w \in \mathbb{R}^D, b \in \mathbb{R}} \sum_{i=1}^n (w \cdot \phi(x_i) + b - y_i)^2 + \lambda (\rho \|w\|_2^2 + (1 - \rho) \|w\|_1), \tag{7}$$

where $\rho \in [0, 1]$ is a constant that balances sparsity characteristics with feature selection ability (Zou and Hastie, 2005) and ϕ is the feature mapping. ρ and ϕ , together with λ , must be tuned during the *Model Selection* phase. Note that Problem (7) is convex and can be solved with any solver, in our case CPLEX (IBM, 2018).

The *Bootstrap* technique (Oneto, 2018) has been exploited as *Model Selection* algorithm, since it is the most effective one in cases like the one described in the paper, when the cardinality of the sample is reasonable (Oneto, 2018). The *Bootstrap* relies on a simple idea: the original dataset \mathcal{D}_n is resampled many (r) times with replacement, to build two independent datasets called learning and validation sets, respectively \mathcal{L}_l^j and \mathcal{V}_v^j , with $j \in \{1, \dots, r\}$. Note that $\mathcal{L}_l^j \cap \mathcal{V}_v^j = \emptyset$. Then, to select the best set of hyperparameters $(\phi, \rho, \lambda)^*$ in the set of possible ones $\{(\phi_1, \rho_1, \lambda_1), (\phi_2, \rho_2, \lambda_2), \dots\}$ for Problem (7) or, in other words, to perform the *Model Selection* phase, the following procedure needs to be applied:

$$(\phi, \rho, \lambda)^* : \arg \min_{\{(\phi_1, \rho_1, \lambda_1), (\phi_2, \rho_2, \lambda_2), \dots\}} \frac{1}{r} \sum_{j=1}^r \hat{L}^{\mathcal{V}_v^j}(h^{\mathcal{L}_l^j}), \tag{8}$$

Time [h]	Mode	Rate of revolution [rpm]	Engine Load [%]
3	Manoeuvre	mostly 500 to 700 rpm	idling and low load
3	Transit	constant near 1000 rpm speed	near maximum load
1	Transit	low speed between 500 and 600 rpm	intermediate loads from 10% to 50%
3	Transit	constant near 1000 rpm speed	near maximum load
4	Manoeuvre	various changing speeds from 500 to 950 rpm	various changing load from 10% to 80%

Table 2: \mathcal{D}_n - Training dataset operational description.

Time [h]	Mode	Rate of revolution	Engine Load
12	Transit	760 rpm	50% to 55%
2	Manoeuvre	various speeds up to 850 rpm	various load from idle to 50%
2	Transit	780 rpm	55% to 60%
1 3/4	Manoeuvre	low 500 to 600 rpm speed	up to 20% load
1 1/2	Transit	400 to 550 rpm	low Load
1/4	Transit	gradually increasing speed from 400 to 780 rpm	gradually increasing load from 10% to 60%
2 1/2	Transit	780 rpm	55% to 60%

Table 3: \mathcal{T}_m - Testing dataset operational description.

where $\hat{L}^{\mathcal{V}_v^j}(h^{\mathcal{L}_l^j})$ is the error on the validation set \mathcal{V}_v^j of the function h^* trained with \mathcal{L}_l^j via Algorithm (7). Note that, in *Bootstrap*, $l = n$ and \mathcal{L}_l^o must be sampled with replacement from \mathcal{D}_n , while $\mathcal{V}_v^o = \mathcal{D}_n \setminus \mathcal{L}_l^o$. The idea behind the *Bootstrap* is simple: $(\phi, \rho, \lambda)^*$ should be the set of hyperparameters which allows the function to work fine on data that has not been used to train the model.

5 Engine Hybrid Model

The proposed *Engine Hybrid Model* is a combination of an *Engine Physical Model* and the *Engine Data-Driven Model*. This requires to modify the *Engine Data-Driven Model* as defined in the previous section in a way to include the mechanistic knowledge of the system. The approach of this paper is a naive one, as the output of the *Engine Physical Model* is used as a new feature that the *Engine Data-Driven Model* can use for training the model.

In the *Engine Hybrid Model* case, the *Engine Physical Model* can be seen as a function of the input x . The *Engine Physical Model*, that we will call h_{EPM} , allows the creation of a new training and testing dataset:

$$\mathcal{D}_n = \left\{ \left(\begin{bmatrix} x_1 \\ h_{\text{EPM}}(x_1) \end{bmatrix}, y_1 \right), \dots, \left(\begin{bmatrix} x_n \\ h_{\text{EPM}}(x_n) \end{bmatrix}, y_n \right) \right\}, \quad \mathcal{T}_m = \left\{ \left(\begin{bmatrix} x_1^t \\ h_{\text{EPM}}(x_1^t) \end{bmatrix}, y_1^t \right), \dots, \left(\begin{bmatrix} x_m^t \\ h_{\text{EPM}}(x_m^t) \end{bmatrix}, y_m^t \right) \right\}$$

Based on these new datasets an *Engine Hybrid Model* can be generated $h_{\text{EHM}} \left([x^T | h_{\text{EPM}}(x)]^T \right)$ and tested. According to this approach, every run of the *Engine Hybrid Model* requires an initial run of the *Engine Physical Model* in order to compute its output $h_{\text{EPM}}(x)$, which allows evaluating $h_{\text{EHM}} \left([x^T | h_{\text{EPM}}(x)]^T \right)$. This is the simplest approach for including new information into the learning process.

6 Results and Discussions

In this section, the results obtained by the *Engine Physical Model*, *Engine Data-Driven Model* and *Engine Hybrid Model*, for the diesel engine temperature prediction, based on the data described in Section 2, are shown and discussed. During the experiments, in the *training phase* the authors exploited one set of training data \mathcal{D}_n , while in the *feed-forward phase* a completely different run of test data \mathcal{T}_m has been used.

During the *Model Selection* the hyperparameter r was set equal to 1000 and the hyperparameter was searched for $\rho \in \{0, 2^{-15}, 2^{-14}, \dots, 2^{-1}, 1-2^{-2}, \dots, 1-2^{-15}, 1\}$ and $\lambda \in \{10^{-6.0}, 10^{-5.5}, \dots, 10^{+6.0}\}$. Finally, for ϕ the following feature mappings $\phi(x) = \{x, [x, x^2], [x, x^2, x^3], \dots, [x, x^2, x^3, \dots, x^{10}]\}$ have been tested.

The two data sets consist of various different manoeuvres using the two control modes described in Geertsma et al. (2017):

- manoeuvre mode: combinator curve with relative low pitch, high engine speed and fast acceleration rates;
- transit mode: combinator curve with higher pitch lower engine speed and slow acceleration rates.

The used training dataset consists of 14 hours of data covering the manoeuvring periods and characteristics reported in Table 2. In Table 3 the description of the used testing dataset, consisting of 21 3/4 hours of data, is reported. The test run was performed utilising the three different modelling approaches to predict the engine temperatures. The results for the turbine exit (T_e) and exhaust receiver temperatures (T_d) are reported from Figure 4 to Figure 9 for the three different numerical models.

Model	T_e - REP [%]	T_d - REP [%]	T_e - RMSE [°C]	T_d - RMSE [°C]
Physical	7.36	10.20	29.7	46.7
Data Driven	2.58	2.67	10.5	12.2
Hybrid Model	2.02	2.14	8.2	9.3

Table 4: Indexes of performances (REP and RMSE) of the models.

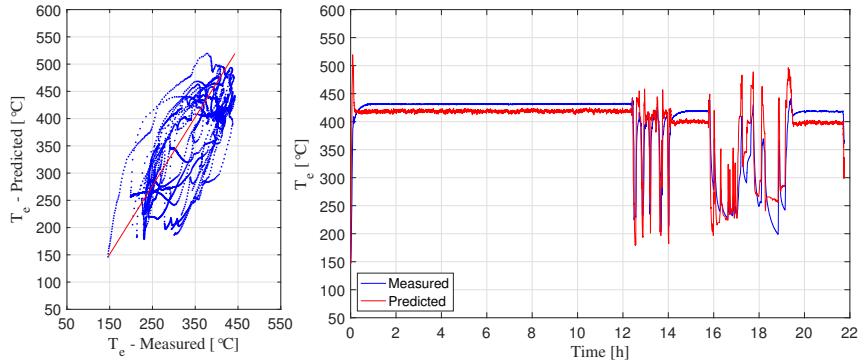


Figure 4: Physical Model results - Exhaust receiver temperature results (T_e).

Finally, Table 4 reports the indexes of performances in terms of REP and the RMSE for the three different models. From the results it is possible to observe that, as expected, the *Hybrid Model* mostly outperform the *Physical Model* and the *Data Driven Model*.

6.1 Physical Model results

The *Engine Physical Model* has been validated using the training data \mathcal{T}_m described in Section 2. The results of the validation are presented in Figure 4 and Figure 5. The thermocouple heat losses and inertia, (1) and (2), were included to better reflect the measured temperatures. The temperature prediction of the *Physical Model* roughly follows the trend of measured values, as shown in Figure 4 and 5. However, the use of one bias and time constant does not accurately represent the dynamic behaviour of the physical heating up and cooling down of the temperature sensor, as the large errors in the scatter plot in Figure 4 and 5 indicate. The work of Jaremkiwicz (2011) and Majdak and Jaremkiwicz (2016) demonstrates an inverse square root relation between the time constant of the thermocouple and the speed of the air flowing past it. Similarly Roberts et al. (2011) describes the heat transfer mechanism that drives the temperature losses; the dependency on flow speed through the *Re*-number, temperature dependent gas properties through the *Pr*-number, the main driving temperature differences, and the radiation heat losses that play an increasingly dominant role as the exhaust gas temperature increases ($q \propto T^4$). While including these effects could further improve the temperature prediction, future work could also focus on establishing these relations with *Machine Learning* techniques as the naive data driven and hybrid approaches already achieve significantly improved results.

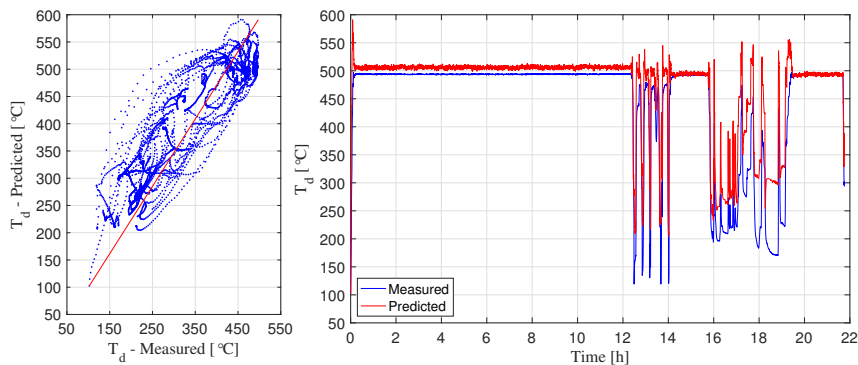
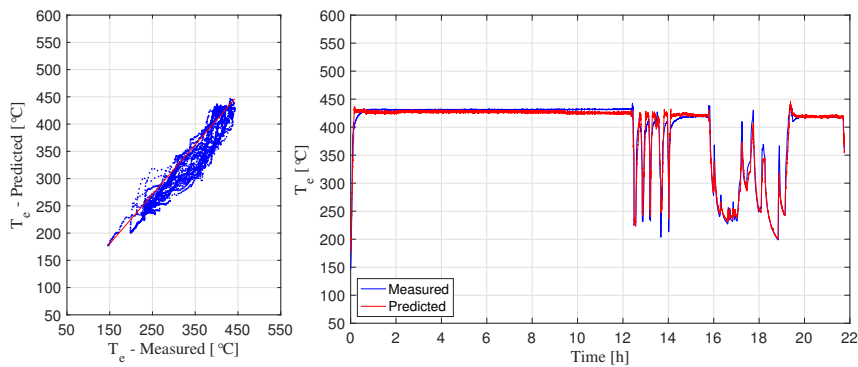
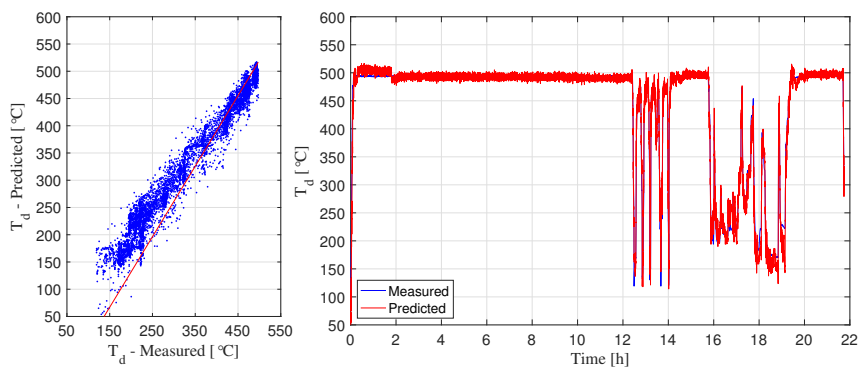


Figure 5: Physical Model results - Turbine exit temperature results (T_d).

Figure 6: Data Driven Model results - Turbine exit temperature results (T_e).Figure 7: Data Driven Model results - Exhaust receiver temperature results (T_d).

6.2 Data Driven Model results

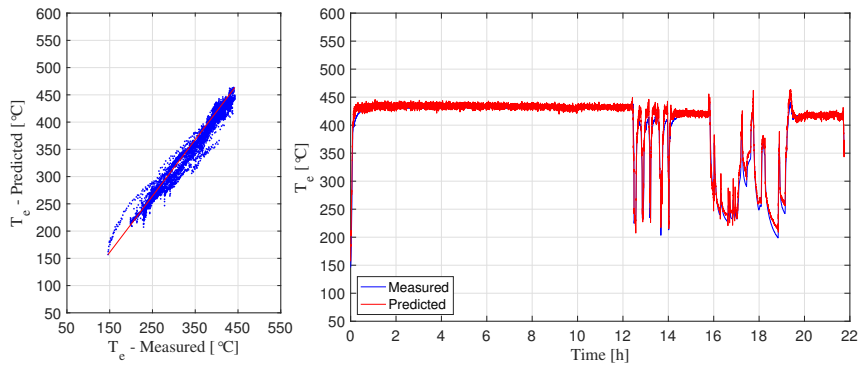
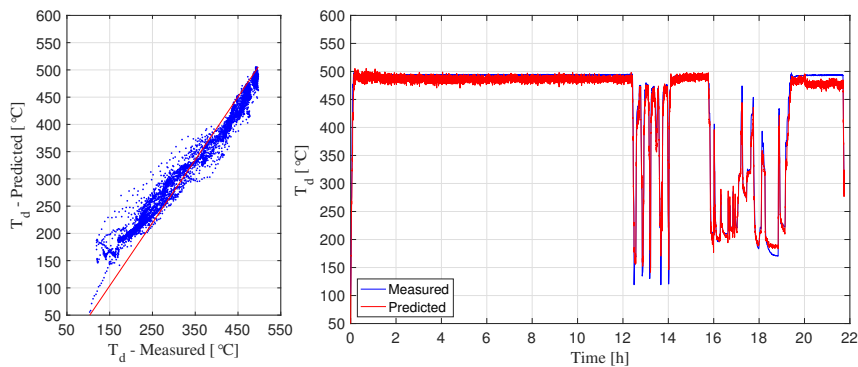
The data driven approach reduces the relative error percentage from 7.36 to 2.58 for the turbine exit temperature and from 10.2 to 2.67 for the exhaust receiver temperature. However, in the constant speed runs the prediction of the exhaust receiver temperature still has a constant error of 10 to 15 °C. This constant error could be caused by the fact that the constant speed runs in the training phase and test phase take place at different engine speeds. The scatter plot reveals that the error increases up to 75 °C during dynamic manoeuvres. While this is a significant improvement compared to the physical model with simple thermal delay, the accuracy might be insufficient for CBM purposes.

6.3 Hybrid Model results

The hybrid model which uses the physical model results in a naïve approach, further reduces the relative error percentage from 2.58 to 2.02 for the turbine exit temperature, and from 2.67 to 2.14 for the exhaust receiver temperature. Moreover, the error in the constant speed runs also reduced further, although the noise on the predicted values increases slightly. The noise is caused by the influence of the fuel pump setpoint which is disturbed by external disturbances such as waves, wind and rudder action. These disturbances do not directly impact temperatures due to the thermal inertia. These inertia effects are not fully addressed in the data driven or hybrid model because the model only uses features at the current time step. Future work could include considering past values of features as well. In addition, a thorough analysis of the data driven and hybrid model could lead to more directed application of data learning techniques, for example by establishing the relationship of the temperature bias and thermal inertia time constant using *Machine Learning* techniques.

6.4 Condition based monitoring framework

The model results presented in this section demonstrate that the hybrid model can be used to predict the engine temperatures during operation in dynamic conditions more accurately. The authors propose to develop a hybrid model further to predict multiple engine parameters and train the hybrid model during operation, to establish trends in the model parameters over time. In particular, the proposed hybrid model can be built during a suitable period when the engine can be considered in healthy condition for a period long enough to observe the system in different operational conditions. Subsequently, the model can be applied and the temperatures computed. The drift in the average behaviour between the predicted and measured parameters can be used as an estimator of the diesel engine

Figure 8: Hybrid Model results - Turbine exit temperature results (T_e).Figure 9: Hybrid Model results - Exhaust receiver temperature results (T_d).

health status, and provide insight about degradation or imminent failures of the engine. For example, increased exhaust receiver temperatures can be used as an indication for exhaust valve or turbine blade degradation or failure the actual temperatures compared to the measured ones. Moreover, the trends in model parameters can also be used as an indication for the deterioration of the engine over time, in particular after collecting and analysing data over a large set of engines and relating them to inspection results of top or complete overhauls. These two methods can then be used to perform condition based maintenance as an alternative for current standard periodic maintenance of diesel engines.

7 Conclusions and further work

Predicting engine temperatures is important to determine wear and failure mechanisms that depend on thermal loading or cause changes in engine operating temperatures. In this paper, the authors have shown that naïve application of data driven models and a hybrid combination of data driven and physical models can improve the accuracy of temperature prediction both during static and dynamic conditions, while existing models do not accurately predict temperatures measured with thermocouples installed on engines in an operational environment under dynamic conditions. A limited amount of operational data was already sufficient to produce reasonably accurate results. Nevertheless, further use of data from the extensive dataset could provide further improvements. Including past measurements or filtering of the results over time could potentially reduce the observed noise on the predicted temperature. Moreover, *Machine Learning* techniques can be used to establish algebraic and differential equations for complex phenomena such as the heating up and cooling down of a temperature sensor in an engine. Subsequently, the hybrid data driven and physical model can be run real-time in parallel with an operational engine to indicate emerging failures early and identify trends that can be used to perform condition based maintenance. These techniques need to be further developed to support future manning reduction for naval vessels.

Acknowledgement

This project is supported by the Royal Netherlands Navy supplying the operational measurement data from one Holland class Patrol Vessel and Damen Schelde Naval Shipbuilding supplying Figure 1.

References

Coraddu, A., Oneto, L., Baldi, F., Anguita, S., 2017. Vessels fuel consumption forecast and trim optimisation: A data analytics perspective. *Ocean Engineering* 130, 351–370.

- Coraddu, A., Oneto, L., Ghio, A., Savio, S., Anguita, D., Figari, M., 2016. Machine learning approaches for improving condition-based maintenance of naval propulsion plants. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment* 230 (1), 136–153.
- Coraddu, A., Oneto, L., Ghio, A., Savio, S., Figari, M., Anguita, D., 2015. Machine learning for wear forecasting of naval assets for condition-based maintenance applications. In: *Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles (ESARS), 2015 International Conference on*.
- Cristianini, N., Shawe-Taylor, J., 2000. *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press.
- Geertsma, R. D., Badon Ghijben, N. A., Zwart, O. R., 2013a. Fire fighting and damage control automation: enabling future crew reduction. In: *Engine as a Weapon International Symposium*.
- Geertsma, R. D., Badon Ghijben, N. A., Zwart, O. R., 2013b. Integration of combat & platform management. In: *Engine as a Weapon International Symposium*.
- Geertsma, R. D., Negenborn, R. R., Visser, K., Loonstijn, M. A., Hopman, J. J., 2017. Pitch control for ships with diesel mechanical and hybrid propulsion: Modelling, validation and performance quantification. *Applied Energy* 206, 1609–1631.
- Grimmelius, H. T., Stapersma, D., 2000. Control optimisation and load prediction for marine diesel engines using a mean value first principle model. Newcastle-upon-Tyne, UK, pp. 212–219.
- Horenberg, S. C., Melaet, A. C. F., 2013. Uniting weapon and marine knowledge. In: *Engine as a Weapon International Symposium*.
- Hountalas, D. T., 2000. Prediction of marine diesel engine performance under fault conditions. Vol. 20. pp. 1753–1783.
- IBM, 2018. User-Manual CPLEX 12.7.1. IBM Software Group.
- Jaremkiewicz, M., 2011. Inverse heat transfer problem encountered in measurement of transient fluid temperature. Publishing of Cracow University of Technology, Cracow.
- Kökkülünk, G., Parlak, A., Erdam, H. H., 2016. Determination of performance degradation of a marine diesel engine by using curve based approach. Vol. 108. pp. 1136–1146.
- Korczewski, Z., 2015. Exhaust gas temperature measurements in diagnostics of turbocharged marine internal combustion engines part i standard measurements. *Polish Maritime Research* 22 (1), 47–54.
- Korczewski, Z., 2016. Exhaust gas temperature measurements in diagnostics of turbocharged marine internal combustion engines part ii dynamic measurements. *Polish Maritime Research* 23 (1), 68–76.
- Leifsson, L., Sævarsdóttir, H., Sigurdhsson, S., Vésteinnsson, A., 2008. Grey-box modeling of an ocean vessel for operational optimization. *Simulation Modelling Practice and Theory* 16 (8), 923–932.
- Majdak, M., Jaremkiewicz, M., 2016. The analysis of thermocouple time constants as a function of fluid velocity. *Measurement Automation Monitoring* 62.
- Nielsen, K. V., Blanke, M., Eriksson, L., 2017. Control-oriented model of molar scavenge oxygen fraction for exhaust recirculation in large diesel engines. Vol. 139. pp. 1–9.
- Oneto, L., 2018. Model selection and error estimation without the agonizing pain. *WIREs Data Mining and Knowledge Discovery*.
- Pedersen, B. P., Larsen, J., 2009. Prediction of full-scale propulsion power using artificial neural networks. In: *International Conference Computer and IT Applications in the Maritime Industries*.
- Petersen, J. P., Jacobsen, D. J., Winther, O., 2012. Statistical modelling for ship propulsion efficiency. *Journal of marine science and technology* 17 (1), 30–39.
- Radonjic, A., Vukadinovic, K., 2015. Application of ensemble neural networks to prediction of towboat shaft power. *Journal of Marine Science and Technology* 20 (1), 64–80.
- Roberts, I., Coney, J., Gibbs, B., 2011. Estimation of radiation losses from sheathed thermocouples. Vol. 31. pp. 2262–2270.
- Rosasco, L., De Vito, E., Caponnetto, A., Piana, M., Verri, A., 2004. Are loss functions all the same? *Neural Computation* 16 (5), 1063–1076.
- Sapra, H., Godjevac, M., Visser, K., Stapersma, D., Dijkstra, C., 2017. Experimental and simulation-based investigations of marine diesel engine performance against static back-pressure. Vol. 204. pp. 78–92.
- Theotokatos, G., Tzelepis, V., 2015. A computational study on the performance and emission parameters mapping of a ship propulsion system. Vol. 229(1). pp. 58–76.
- Tikhonov, A. N., Arsenin, V. Y., 1979. *Methods for solving ill-posed problems*. Nauka, Moscow.
- Vapnik, V. N., 1998. *Statistical learning theory*. Wiley.
- Wolpert, D. H., Macready, W. G., 1997. No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation* 1 (1), 67–82.
- Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2), 301–320.