Marine engines combustion diagnostics employing fourier series and ANN

Chaitanya Patil*[1], Gerasimos Theotokatos[1], and Konstantinos Milioulis[2]
[1]Maritime Safety Research Centre, University of Strathclyde, UK
[2]Datum Electronics, UK

Abstract

Safe operations of marine engines is ensured by appropriate maintenance techniques requiring accurate assessment of engine's health status. The use of machine learning methods can considerably enhance the combustion diagnostics and hence facilitate the cost-effective and timely maintenance of marine engines. This study aims at assessing the potential of Fourier series coefficients (FC) obtained from in cylinder pressure signal and developing an artificial neural network (ANN) model that can support the engine diagnostics of marine engines. A ferry ship with two propulsion engines of the four-stroke type was employed as the reference system in this study. Digital twin of the thermodynamic zero dimensional type, which was calibrated by using the engines shop test measurements, is employed to generate the required data-sets in the whole engine envelop, whilst considering the most typical engine anomalies, including degradation and faults. The results demonstrate that first 20 harmonics contains required information to estimate fault severity within 0.016 RMSE range.

Introduction

Engine health diagnosis is important to ensure the uninterrupted functioning of marine engines. Literature depicts various efforts taken in the field of engine diagnostics. Studies have shown that traditional fault diagnosis methods, such as visual inspections and manual testing, are time-consuming and often lead to incorrect results. In recent years, Artificial Intelligence (AI) and Machine Learning (ML) algorithms have been developed to analyse large amounts of data from marine engines to detect faults in real-time. Xi et al. [12] developed a classifier to avoid human errors using independant component analysis (ICA) for vibration data obtained from the various engine locations. Tsaganos et al. [6] proposed an ensemble (multi technique) model to classify the common occurring faults in marine two-stroke engines. Extreme learning ensemble is developed by Kowalski et al. [2] to classify around 14 faults in four stroke diesel engines. In brief, ML is widely used for performance estimation, diagnosis and prognosis of combustion engines. The data driven models based on ML has proven to be cost effective solutions to facilitate the maintenance process of marine engines.

The change in thermodynamic/mechanical/physical settings of marine engines are reflected on instantaneous in-cylinder pressure of marine engines. Moreover, efforts have been taken to extract the combustion information from the in-cylinder pressure curve. This signal is important to sense anomalies within the engine working envelope. Fourier analysis is widely used to convert the periodic signals into frequency domain. The harmonics orders in terms of Fourier coefficients (FC) have potential to extract relevant information from the pressure signal, and can be further used for diagnosis instead of the actual pressure signal.

This study aims at assessing the potential of using FC obtained from in cylinder pressure signal for fault diagnosis that can support the engine diagnostics of marine engines. Fourier analysis is used to convert the pressure signals into FC corresponding to different harmonics orders. The impact of faults on these FC are analysed with variance analysis to select N potential harmonics carrying appropriate information for fault diagnosis. The effect on number of harmonics selected for diagnosis is also assessed through RMSE on validation data. Finally, effective model with confirmed harmonics N is tested on test data set and the performance is quantified on the basis of RMSE and training time.

Methodology

The methodological approach consists of the following six steps:

Step 1 Data generation – A validated digital twin of the investigated marine engine is employed to generate *healthy* and *faulty* data-sets.

Step 2 Fourier analysis – Conversion of in-cylinder pressure signals to discrete FC.

Step 3 Data standardisation and splitting – The data (input:engine speed, power, FC and output:fault severity) is standardised by centering to their mean values. The generated data from Step 1 is split in 3 parts called training, validation and test data sets.

Step 4 Model setup and training – Artificial Neural Networks (ANN) are set up based on the selected input and output parameters.

Step 5 Model testing – The trained model is tested on the test data (separated in step 4).

Data generation

The engine used for this study is the four-stroke Wärtsilä 9L46C marine engine, which is a nine cylinder turbocharged medium speed engine. The specifica-

*Corresponding author: chaitanya.patil@strath.ac.uk

tions of the engine are described in Table 1. The employed digital twin (DT) is of the zero-dimensional type and uses semi-phenomenological models and widely acknowledged formulae to represent the engine processes. The detailed description of the employed digital twin is provided in [7, 8, 10].

Maximum Continuous Rating point	9,450 kW @ 500 RPM
No. of Cylinders	9
Cylinder Bore	460 mm
Clutch-in Speed	300 RPM
Turbocharger	ABB TPL 77-A30

Table 1: Investigated marine engine specifications. [11]

The zero-dimensional thermodynamic DT for the investigated engine is calibrated and validated by considering the steady state measured parameters from the engine shop trials (factory acceptance tests). The employed DT calibration and validation processes for the investigated marine engine are described in [7–9]. This DT was further validated against in-cylinder pressure measurements acquired during the normal ship operation at five steady state operating points.

Subsequently, the validated DT is employed to generate the required data sets considering the investigated engine operation at both *healthy* and *faulty* conditions; the latter are associated with several engine components faults and degradation. The four most frequent faults/degradation are: fuel injection issues, increased friction losses, blowby, harge Air Cooler (CAC) fouling.

The faults and their severity β_{fault} range considered for the simulations are listed in Table 2. It should be noted that this study considers faults/degradation for one cylinder of a multi-cylinder engine and can be extended to other cylinders. The cases of different combination of faults/degradation on cylinders is out of the scope of this study.

Table 2: Considered faults/degradation and input parameters specification in the DT.

Faults/degradation Description	Parameter Calculation	Range of β_i
Injection advance	$\theta_{SOI,\text{healthy}}(1-\beta_{SOI})$	0–0.6
Engine friction losses increase	$fmep_{\text{healthy}}(1+\beta_{FMEP})$	0–0.6
Blowby	$A_{BB,\text{healthy}}\beta_{Blowby}$	0–0.6
Charge Air Cooler (CAC) Fouling	$\eta_{AC,\text{healthy}}(1-\beta_{CAC})$	0–0.6

To facilitate the fault/degradation analysis over the whole engine operating envelope, 50 operating points corresponding to steady state conditions were considered.

Fourier series coefficients (FC)

The number of points from the pressure signal can be reduced by sampling it with lower sampling rate. However, the risk of loosing information renders this method ineffective. Fourier analysis [4] is a method to represent continuous periodic signals through superposition of harmonically related sine and cosine signals. Several efforts to reconstruct the in-cylinder pressure signals using FC were reported in the pertinent literature [1, 5, 13]. Moreover, FC are proven to be useful to reduce the dimensionality in other machine learning applications as reported in [3]. This method has a potential to represent a typical pressure signal with less than 100 points (coefficients).

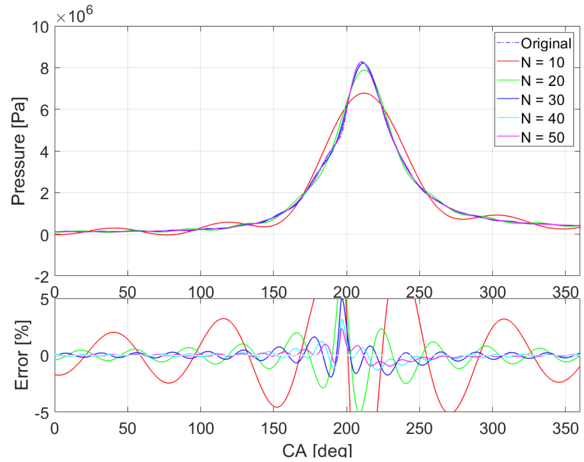


Figure 1: Reconstruction of pressure signal (first 360°CA) from N number initial harmonics using 2N+1 FC.

On the basis of this concept, the derived data sets of the in-cylinder pressure, which are functions of the crank angle corresponding to the time domain, as the crank angle is a function of time (for steady state conditions, ϕ(°CA)= 6 $N(rpm)t(s)$), can be converted to the frequency domain by the use of discrete coefficients $(C_1, C_2, ..C_N)$, according to the following equation:

$$P(\phi) \Rightarrow f(C_1, C_2, ...) \qquad (1)$$

The in-cylinder pressure signal of four stroke engine cycle has period T equals to 720°CA, and can be converted into Fourier series using the equations reported in [13].

The total number of harmonic orders required to exactly represent the pressure signal (resulting in zero error) tends to infinity. Figure 1 shows the reconstructed in-cylinder pressure profile considering N numbers of initial harmonic orders considered from 2N+1 FC. As N increases, the error in reconstruction of original pressure signal reduces. Higher values of N will be required to accurately reconstruct the pressure signal. However, the diagnostic model demands only required amount of information which is helpful for fault prediction. Typically the lower harmonic orders convey most of the required information, thus, the in-cylinder pressure signals can be represented by using a lower number of input parameters resulting in dimensions reduction as required in developing data-driven models. Therefore, corresponding value of N is selected by variance analysis presented in results section to provide appropriate number of coefficients as input to the diagnostic model.

Data standardisation and splitting

Scaling of parameters allows the ML algorithms to focus on the relationships between the parameters, rather than the scale of the values. All numerical attributes (input and output parameters of the developed models), which include engine performance parameters and FC along with the faults severity parameters (β_{fault}), are standardised by removing their mean and scaling them to the unit variance.

The populated data from simulations in step 1 is split in three parts for training, validation and test data sets. 15% of the total data is used for testing, while remaining dat is used for training purpose. Another split of training data is carried out by taking 15% of the training data as validation data to validate the ANN during training phase.

Model setup and training

The ANN is developed with input [engine speed, engine power, FC (considering N harmonics)] and output [β_{CAC},β_{FMEP},β_{SOI},β_{blowby}] using *Keras* and *TensorFlow* package.

MultiLayer Perceptron (MLP) networks are type of ANN, which are also known as multilayer feed forward networks, are widely used in practical applications. Figure 2 showcases the structure of ANN used in this study. Two hidden layers with 10 nodes for each hidden layer

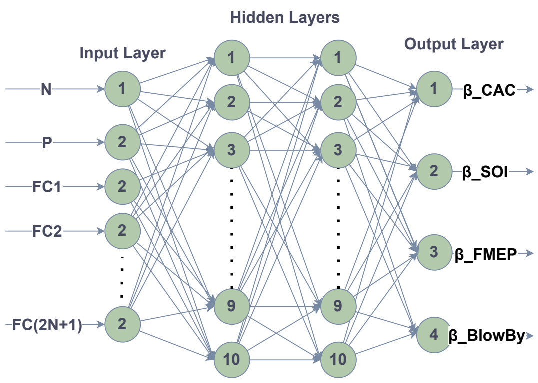
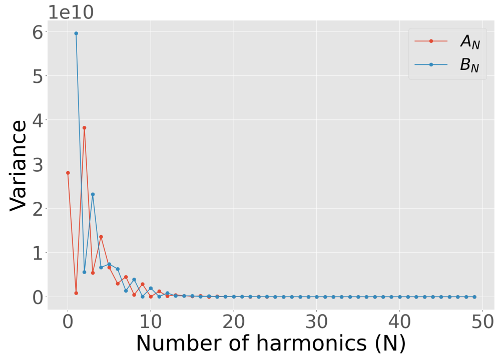


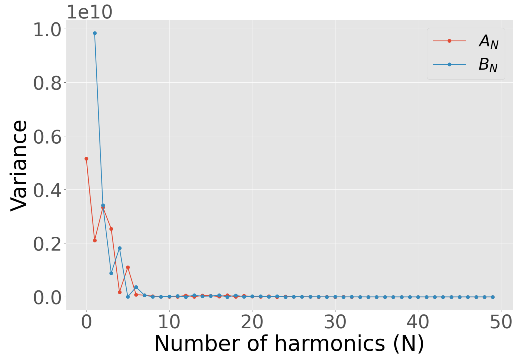
Figure 2: Artificial Neural Network with 2 hidden layers estimating fault severity of most frequent faults

of ANN are selected for this study. Considering the non-linear relationship between inputs and output, an activation function called Exponential Linear Unit (ELU) is used at each neuron of hidden layer followed by a linear activation function at the output layer.

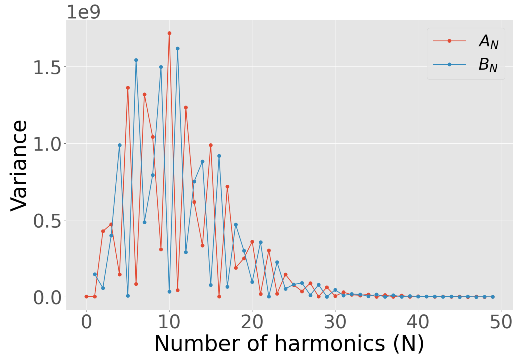
Metric selection is necessary to monitor the performance of the model during training as well as validation on test data. The developed model estimates fault severity index as a numerical value. The Root Mean Square Error is selected to monitor the performance of the model. The training process involves, minimisation of RMSE between true and predicted value β_{fault} by the model. The *Adam* optimiser is used with adapting learning rate to minimise the monitored RMSE. Early stopping constraint is introduced monitoring consecutive it-



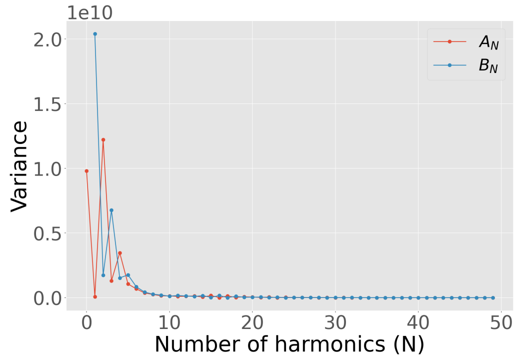
(a) CAC Fouling



(b) Friction loss



(c) Injection advance



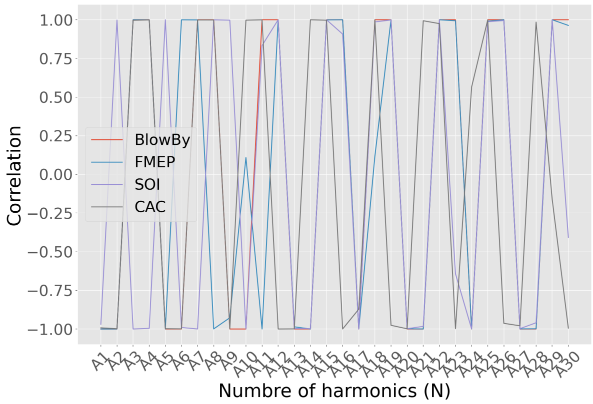
(d) BlowBy

Figure 3: Variance of A_N and B_N coefficients for $N \in [1,2,..,N]$ for each fault (a,b,c,d) present alone at 462.5 RPM and full load.

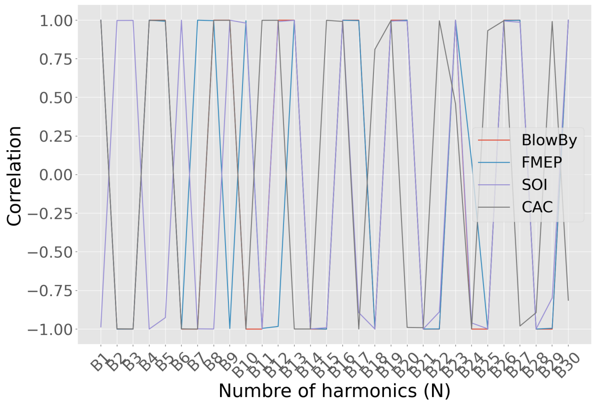
eration losses to check if the error is stabilised for stopping the training process earlier saving training time.

Model testing

The trained model is further tested on the test data set separated in step 4. The performance of the model is always verified on the test data as the model has never seen this data in training period and makes the evaluation unbiased. The test results are presented in next section.



(a) A_N



(b) B_N

Figure 4: Correlation of A_N and B_N coefficients for $N \in [1,2,..,N]$ with β_{fault} present alone at 462.5 RPM and full load.

Results

This section presents the results of assessing FC to be used for a fault diagnosis.

Every fault with severity (β_{fault}) has an impact on the harmonics of pressure signal calculated from step 2. To verify this impact, variance of FC (A_N and B_N) considering N^{th} harmonics over the β_{fault} is plotted for each fault separately in Figure 3. The impact shows high variance on initial harmonics ($N \leq 20$) for β_{FMEP},β_{BlowBy} and β_{CAC}. While, fault due to injection advance reflects on higher order harmonics ($N \leq 30$). Moreover, The correlation of A_N and B_N up to 30^{th} harmonic is plotted with β_{fault} for each fault in Figure 4. It shows all the coefficients are highly correlated with ± 1 value. Therefore, it is evident that, the initial 20 fundamental harmonics of the pressure signal are crucial information carriers for fault diagnosis.

The structure of ANN using 20 harmonics (41 FC) with 2 hidden layers and 4 output nodes for each fault severity in presented in Table 3.

Layer (type)	Output shape	Parameters
Input layer	(None,43)	0
Hidden layer (Dense)	(None,10)	440
Hidden Layer (Dense)	(None,10)	110
Output layer	(None,4)	44

Table 3: Structure of ANN using N=20 harmonics as input along with engine speed and power to predict β_{fault} of four faults

The RMSE on the validation data (separated in step 4) during training over 200 epochs is presented in Figure 5. The error reduces as the ANN is trained with increasing number of times (epochs) the data is presented to the model. The learning rate changes due to *Adam* optimiser speeding the training process the reduce the RMSE.

various ANN architecture with different number of harmonics (N) are considered to verify above variance analysis to assess the potential of N number of harmonics to diagnose the faults. The RMSE surely reduces with higher harmonics however, it gets saturated after N=20. The RMSE of ANN with 20 harmonics, after training with 174 epochs is 0.0162. The training time increases with increased number input to ANN with number of harmonics. Total training time for the model created with 20 harmonics is 91.3 seconds. Thus, only initial 20 harmonics are enough to predict the severity of the frequent faults with good accuracy and training time.

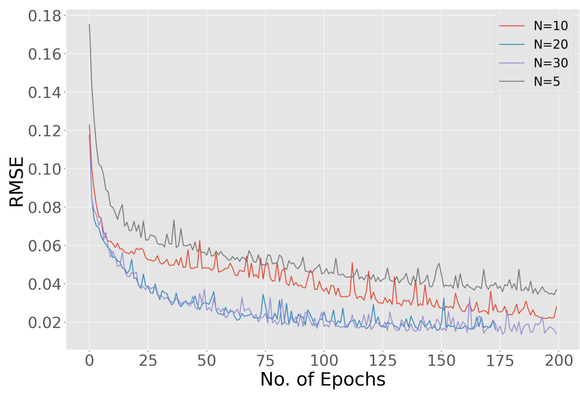


Figure 5: RMSE on the validation data set during training phase of the ANN

The trained model is further tested on the test data separated in the step 4. As this data was kept aside from the model during training period, the model is not biased for this data set. The RMSE on this data set to predict the β_{fault} for each fault is presented in the Table 4. The overall root means square error for the ANN with 20 initial harmonics is 0.016 on the test data.

Fault	RMSE
BlowBy leakage	0.012
Friction	0.011
Start of injection	0.0096
Cooler fouling	0.025

Table 4: RMSE to predict the β_{fault} using the ANN diagnostic model.

Conclusions

The assessment of FC to predict the faults in marine engines is carried out successfully, highlighting the potential of initial harmonics for estimating the most frequent fault severity. The corresponding FC (2N+1) for N harmonics calculated from measured in-cylinder pressure signal are used as inputs to the ANN along with engine speed and power. The fault severity (β_{fault}) of four combustion related faults is estimated by the ANN. The principle findings of this study are,

1. The impact on FC due to faulty operating conditions, reduces with higher orders of harmonics
2. The variance analysis implies that initial 20 fundamental harmonics of the pressure signal are crucial information carriers for fault diagnosis.
3. The saturating RMSE of ANN considering harmonics more than 20 confirms that, N=20 is enough to predict the severity of the frequent faults with good accuracy and training time (93s).
4. The overall RMSE of the ANN diagnostic model to predict β_{fault}, considering 20 initial harmonics on the test data is 0.016

Acknowledgement

This study was carried out in the framework of the i-HEATS project funded by the Innovate UK Smart Grants (agreement No 99958). The authors affiliated with the MSRC also greatly acknowledge the funding from DNV AS and RCCL for the MSRC establishment and operation. The opinions expressed herein are those of the authors and should not be construed to reflect the views of Innovate UK, DNV AS and RCCL.

References

[1] Roger Johnsson. Cylinder pressure reconstruction based on complex radial basis function networks from vibration and speed signals. *Mechanical Systems and Signal Processing*, 20:1923–1940, 11 2006.

[2] Jerzy Kowalski, Bartosz Krawczyk, and Michał Woźniak. Fault diagnosis of marine 4-stroke diesel engines using a one-vs-one extreme learning ensemble. *Engineering Applications of Artificial Intelligence*, 57:134–141, 2017.

[3] V.N. Manjunath Aradhya, G. Hemantha Kumar, and S. Noushath. Multilingual ocr system for south indian scripts and english documents: An approach based on fourier transform and principal component analysis. *Engineering Applications of Artificial Intelligence*, 21(4):658–668, 2008.

[4] Hagit Shatkay. The fourier transform-a primer. *Brown University*, pages 1–20, 1995.

[5] Dinu Taraza, Naeim A Henein, Mangesh J Gade, and Walter Bryzik. Cylinder pressure reconstruction from crankshaft speed measurement in a four-stroke single cylinder diesel engine. In *ASME Internal Combustion Engine Division Spring Technical Conference*, pages 387–395, 04 2005.

[6] G Tsaganos, N Nikitakos, D Dalaklis, AI Ölcer, and D Papachristos. Machine learning algorithms in shipping: improving engine fault detection and diagnosis via ensemble methods. *WMU Journal of Maritime Affairs*, 19(1):51–72, 2020.

[7] K M Tsitsilonis, Gerasimos Theotokatos, and A Coventry. Development of an intelligent marine engines health assessment system based on digital-twins and data-driven models. In *Annual Marine Technology Conference*, pages 1–12, 2021.

[8] Konstantinos Tsitsilonis and Gerasimos Theotokatos. Engine malfunctioning conditions identification through instantaneous crankshaft torque measurement analysis. *Applied Sciences*, 11(8):3522, 2021.

[9] Konstantinos-Marios Tsitsilonis and Gerasimos Theotokatos. A novel method for in-cylinder pressure prediction using the engine instantaneous crankshaft torque. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 236(1):131–149, 2022.

[10] Konstantinos-Marios Tsitsilonis, Gerasimos Theotokatos, Chaitanya Patil, and Andrea Coraddu. Health assessment framework of marine engines enabled by digital twins. *International Journal of Engine Research*, page 14680874221146835, 2022.

[11] Wartsila. Project guide for marine applications: Wartsila 46c, 2001. Accessed: 2022-10-06.

[12] Wenkui Xi, Zhixiong Li, Zhe Tian, and Zhihe Duan. A feature extraction and visualization method for fault detection of marine diesel engines. *Measurement: Journal of the International Measurement Confederation*, 116:429–437, 2 2018.

[13] Pin Zeng and Dennis N Assanis. Cylinder pressure reconstruction and its application to heat transfer analysis. *SAE paper*, page 0922, 2004.