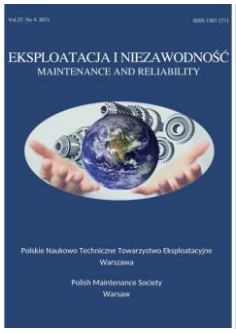


Eksploatacja i Niezawodnosc – Maintenance and Reliability
Volume 25 (2023), Issue 4
journal homepage: http://www.ein.org.pl

Article citation info:
Žvirblis T, Hunicz J, Matijošius J, Rimkus A, Kilikevičius A, Gęca M, Improving Diesel Engine Reliability Using an Optimal Prognostic Model to Predict Diesel Engine Emissions and Performance Using Pure Diesel and Hydrogenated Vegetable Oil, Eksploatacja i Niezawodnosc – Maintenance and Reliability 2023: 25(4) http://doi.org/10.17531/ein/174358

Improving Diesel Engine Reliability Using an Optimal Prognostic Model to Predict Diesel Engine Emissions and Performance Using Pure Diesel and Hydrogenated Vegetable Oil

Indexed by:
Web of Science Group

Tadas Žvirblis[a], Jacek Hunicz[b], Jonas Matijošius[c,*], Alfredas Rimkus[d], Artūras Kilikevičius[c], Michał Gęca[b]

[a] Institute of Data Science and Digital Technologies, Vilnius University, Lithuania
[b] Faculty of Mechanical Engineering, Lublin University of Technology, Poland
[c] Institute of Mechanical science, Faculty of Mechanical Engineering, Vilnius Gediminas Technical University, Lithuania
[d] Department of Automobile Engineering, Vilnius Gediminas Technical University, Lithuania

Highlights

- A three-step algorithm based on statistical prognostic models was used to implement diesel engine reliability improvements.
- The creation of a prolific and effective ANCOVA prognostic model.
- ANCOVA was exceedingly accurate in predicting 95% of the studied parameters.

Abstract

The reliability of internal combustion engines becomes an important aspect when traditional fuels with biofuels. Therefore, the development of prognostic models becomes very important for evaluating and predicting the replacement of traditional fuels with biofuels in internal combustion engines. The models have been made to model AVL 5402 engine emission, vibration, and sound pressure parameters using a three-stage statistical regression models. The fifteen parameters might be accurately predicted by a single statistic presented here. Both fuel type (diesel fuel and HVO) and engine parameters that can be adjusted were considered, since this analysis followed the symmetry of the methods. The data analysis process included three distinct steps and symmetric statistical regression testing was performed. The algorithm examined the effectiveness of various engine settings. Finally, the optimal fixed engine parameter and the optimal statistic were used to construct an ANCOVA model. The ANCOVA model improved the accuracy of prediction for all fifteen missing parameters.

Keywords
engine's reliability, statistical regression analysis; linear regression models; ANCOVA; MAPE; hydrotreated vegetable oil

1. Introduction

Process reliability reduces process internal combustion engine's failures. Process reliability is defined, evaluated, and is very important to prognoze the internal combustion engine's performance [27]. In addition to this, it makes it possible to acquire reliability estimates as well as a hierarchy [6]. Another aspect is the reliability of the internal combustion engine when using biofuels. The main advantage of such fuels is lower exhaust gas emissions [42]. Therefore, when evaluating the reliability of engines, it is necessary to take into account not only engine operating parameters, but also exhaust emissions.

(*) Corresponding author.
E-mail addresses: T. Žvirblis (ORCID: 0000-0002-9368-8726) tadas.zvirblis@mf.vu.lt, J. Hunicz (ORCID: 0000-0002-0188-6419) j.hunicz@pollub.pl, J. Matijošius (ORCID: 0000-0001-6006-9470) jonas.matijosius@vgtu.lt, A. Rimkus (ORCID: 0000-0001-8995-7180) alfredas.rimkus@vilniustech.lt, A. Kilikevičius (ORCID: 0000-0002-4074-1773) arturas.kilikevicius@vilniustech.lt, M. Gęca (ORCID: 0000-0002-2226-1645) michal.geca@pollub.pl

Requirements for engine emissions are becoming increasingly stringent for environmental benefits [47]. Decarbonization programs have been implemented for using clean, low-carbon fuels [37]. Ambitious goals to reduce the concentrations of hazardous compounds in the exhaust gases were supplemented by strict requirements for the reduction of the concentration and volume of solid particles [16], which have already been transferred to the legal basis (requirements of Euro Standard 6) [15]. A new technique of measurement employing random measurement sections that achieves these objectives. The proposed method corresponds to the random nature of road traffic situations while maintaining the similarity of hazardous exhaust component emissions [9].

The harmfulness of exhaust gases is limited by using alternative fuels [4]. The engine runs noisier and performs worse as the amount of ethanol additive in the fuel mixture is increased. This problem has been addressed using a variety of exhaust gas regulation (EGR) situations, where increasing the EGR rate led to a reduction in thermal efficiency [33]. Another method is the supply of hydrogen, but the simultaneous supply of ethanol and hydrogen increased the concentration of soot [13, 44]. The usage of oxygen-diesel mixtures is an efficient technique to minimize soot concentrations because it allows the harmfulness of exhaust gas components to be reduced [19] (except for NO_X, whose concentration increases at higher temperatures [28]. These can be various dibutyl maleate and diesel [34] and biodiesel mixtures (starting with palm oil [29], fats [5], rapeseed oil [31] and ending with diesel esters. Another option is to use hydrogenated vegetable oil (HVO) blends with diesel and soybean oil methyl ester with diesel. The qualities of HVO mixtures with diesel not only improve the environmental indicators of the engine, but also allow for improved performance characteristics [1, 45], especially by optimizing the main fuel injection time [12]. The environmental impact of soybean oil methyl ester mixes with diesel has also been emphasized, highlighting the stability and tarability of such blends [32]. Another key issue is the viscosity of soybean oil methyl ester, which is ten times that of diesel, and its use is restricted, particularly at subzero ambient temperatures [2].

Various machine learning algorithms, including artificial neural networks, relevance vector method, support vector machine, genetic algorithm, response surface method, and gene expression programming, are extensively employed in the prediction of internal combustion engine performance [20]. A novel approach for assessing the dependability of aeroengine cooling blades by employing a multivariate ensembles-based hierarchical linking strategy there was presented. The proposed methodology aims to improve the accuracy and efficiency of computations by developing machine learning models [25]. The DLR-SS method, which is a deep learning approach, successfully addresses the challenges posed by high nonlinearity and correlated interactions in the probabilistic evaluation of CCF damage. This is achieved by the utilization of fatigue life models and synchronous mapping-based models [23].

Over the past decades, researchers have introduced a number of prognostic models for engine emissions, one of the first comprehensive reviews of the models [10]. Predictive maintenance improves the availability, cost, and dependability of high-value assets. A novel prognostic function monitors diesel fuel injector deterioration. The feature value is determined using the thermodynamics of the engine and data from the engine controller to provide a simple and accurate two-stage hybrid linear regression model for the temperature of the exhaust gas [40]. When compared to the temperature that is produced by the statinary heating device, the flame temperature that is produced by the combustion engine is on average two times higher [38]. Using a new engine sensor and an exhaust gas temperature sensor, an artificial neural network identifies knocks. Turbocharged gasoline-powered automobile test travels yield 5 million data sets. Filtered data reduce network errors [17]. The Six-Sigma method can be applied to numerical data that violates the assumption of normal distribution. The power transformation is proposed using multiple criteria decision analysis (MCDA) for the asymmetry coefficient and kurtosis coefficient. To test distribution normality, the Jarque-Bera statistic was minimised [7]. A method of using statistical models to investigate and predict the outcomes of ongoing processes. In some cases, the available classical approaches cannot produce information of a sufficient level of reliability [8]. Single-LSTM (long-short-term memory) models predict gas-path performance for each aero-engine. Operating conditions and time series are modelled together. SALNT (LSTM neural tree), a new prognostic model, incorporates decision trees and LSTM.

SALNT helps the sample find its best predicted gas-path performance model [26]. Gaussian process regression predicted engine performance and exhaust emissions. Bayesian hyperparameter optimisation improved predictive model training. Prophetic models successfully anticipated engine performance and exhaust emissions. Bayesian optimised Gaussian process regression creates a reliable engine performance and emission prognostic model [3].

ANOVA models are most commonly used for predicting how engine emission levels depend on fixed engine parameters such as fuel type, engine power and loads, rpm, etc. ANOVA model is a simple and convenient statistical method when it is necessary to evaluate various individual qualitative parameters and their interactions to predict engine emissions. Sivaramakrishnan and Ravikumar [36] conducted ANOVA models for brake thermal efficiency, brake specific fuel consumption, HC, CO, and NO_X. Compression ratio, fuel mixtures and engine power were selected as independent variables. The adapted R^2 of the realized models was extremely high and reached >90%, while the adapted R^2 of brake thermal efficiency, specific brake fuel consumption and nitrogen oxides were equal to 100%. In the conclusions, the authors additionally note that the experimental design including statistical analysis accelerated the data analysis and reduced the number of experiments. Ramachander et al. [30] publication contains prognostic models based on ANOVA methodology. For more than ten different engine parameters, prognostic models were built with engine load, fuel injection pressure and time input parameters. The constructed ANOVA model showed that engine load was statistically significant for almost all predicted parameters (only NO_X was non-significant). Fuel injection pressure and timing as prognostic factors were significant for a smaller number of predicted parameters. When creating a prognostic model, the authors additionally included pairwise interactions of all input parameters and quadratic expressions of the input parameters themselves. The values of the model parameter estimates indicated that not all interactions are affected by the model output. In this case, it is recommended that non-significant interaction factors should be removed from the model, but it is not clear in the publication whether this was done by the authors. Two-factor ANOVA model was investigated to evaluate the biodiesel blends that would generate the lowest vibrations [39]. The output parameter of the model was the RMS of the vibration signal, and the input parameters (or factors) were two: fuel type (9 variants of the biodiesel mixture) and engine speed (7 variants). Vibration data was obtained for two engine conditions, before and after technical inspection. The ANOVA model showed that regardless of whether the engine was pre- or post-maintenance, fuel type, rpm and the interaction of both of these factors were statistically significant in predicting vibration RMS values ($p < 0.001$). The study further revealed that the highest RMS values of vibration are achieved at 1800 and 2000 rpm. for engine speed. It is also important to note that the RMS of the vibration signal was reduced by 12% after the engine was inspected.

A similar study was conducted to evaluate the biodiesel blends that would generate the lowest vibrations by Uludamar et al. [41]. Linear and non-linear regression models for predicting engine vibration and sound pressure were presented. During the experiment, three types of fuel mixtures (low-sulphur diesel mixed with sunflower, rapeseed and corn oils) were used, and the engine speed was varied (1200, 1500, 1800, 2100 and 2400 rpm). According to the authors, the root mean square deviation is the most important measure of the vibration and sound pressure value, so it was used as the output parameter of the model. The input parameters of the predictive models were fuel density, ketone number, viscosity, and minimum heating value. The goodness-of-fit of the constructed models was assessed using the mean absolute percentage error (MAPE). The RMS values of the vibration signal were found to be slightly worse predicted than the RMS of the sound pressure, but overall the models were accurate and MAPE was less than 1% in all cases.

Khurana et al. [21] reviewed the recent research performed since last decade (2008 to 2020) which investigated machine learning approaches for modeling engine emissions. Authors identified that during research period these advanced techniques as K-nearest neighbor, Bagged Decision Trees, Neural Network, Support Vector Machine, CHAID model, adaptive neuro-fuzzy approach, back-propagation with Levenberg–Marquardt algorithm, and other ANN were the most popular for predicting of engine emissions. A double-layer perceptron neural network for predicting as many as 12 output parameters was compiled and described by Hosseini et al. [18]. The input parameters were

fuel mixture, engine speed, ambient air pressure, relative humidity, oil temperature, exhaust oxygen content, fuel consumption, intake manifold pressure, lower heating value, exhaust gas temperature, fuel density and fuel viscosity. The output parameters were: engine performance (power and torque), emissions (CO, CO_2, UHC and NOx) and vibration. The R^2 for training, validation and testing of the constructed model was as high as 99.99%, 99.94% and 99.95%, respectively. The authors conclude that the constructed two-layer perceptron is a powerful tool for predicting engine performance, emissions, and vibration characteristics.

Several other machine learning models were investigated by Zhang et al. [43]. Random forest, support vector regression, and artificial neural networks were used for prediction of engine efficiency and emission performance. Authors conclude that machine learning models can help to predict engine performance and emissions, however, it required heavy tuning of the hyperparameters, such as the net structure.

Direct injection diesel engines were improved ANN (Artificial Neural Network) to estimate emissions and performance. For some parameters, the model agreed with the experimental data on engine performance and emissions [11]. Using experimental data and a multilayer perceptron architecture, an ANN-based prediction framework for engine functioning and exhaust parameters was developed. 5% of the experimental predictions were inaccurate [35].

An increasing number of publications on this topic shows that prediction of engine performance and emission is not completely solved. According to a survey of the literature, developing statistical prediction models for exhaust emission of ignition engines is still difficult. As prognostic models have continued to evolve, accuracy results have reached little improvement, moreover the complexity of algorithms has grown exponentially. It has been observed that researchers hardly use vibration data in the development of predictive models and rely only on fixed engine parameters.

During this study, vibration and sound pressure data were included in the developing of prognostic model of engine performance and emissions. Several ANOVA prognostic models were developed using statistical optimization process. RMS values of vibration and sound pressure data in combination with injection and EGR ratio were used as prediction model input parameters. The primary objective of this study was to investigate the optimal prediction model for engine performance and emissions parameters dependent from vibration, sound pressure changing injection timing, and EGR rate.

2. Methodology and Data Description

2.1. Experimental Engine

The test object was a single-cylinder, four-stroke AVL type 5402 CR DI compression ignition (CI) engine. Detailed specifications of the test engine are given in Table 1.

Table 1. Research engine AVL 5402 specifications.

Displacement	510.5 cm^3	
Compression ratio	17:1	
Bore / Stroke	85 mm / 90 mm	
No. of valves	4	
Combustion type	Direct injection	
Max. fuel pressure	180 MPa	
Injection system	Common Rail CP4.1	
Engine management	AVL-RPEMS, ETK7-Bosch	
Valve	**open**	**close**
Intake	712 °CA	226 °CA
Exhaust	488 °CA	18 °CA

The displacement volume was 510 cm^3 and the compression ratio was 17:1. The engine had an internal combustion chamber with four valves. The valves were inclined at an angle of 3.5 °

Fuel was supplied by a seven-hole electromagnetic injector and a Bosch CP4.1 high-pressure fuel system. Injection parameters were controlled by a fully open Bosch engine

controller and Etas INCA software. The tests were performed at med-load regime and the engine was run as naturally aspirated. The amount of exhaust gas from the recirculation was regulated by a butterfly valve. In addition, an exhaust gas cooler was installed on the bench to keep the temperature of EGR supplied to the intake manifold constant. To enable high EGR rates, a backpressure valve was installed on the exhaust pipe.

2.2. Engine Testing System

The research engine was mounted on the test stand equipped with an asynchronous motor dynamometer suitable for AVL single cylinder research engines series 540. The AVL 733S dynamic fuel meter was used to measure fuel consumption. The fuel temperature was obtained with an AVL 753C fuel temperature conditioner. In addition, sensors TP-361 and TP-204 were used to measure air, exhaust gas, cooling liquid and lubricating oil temperatures.

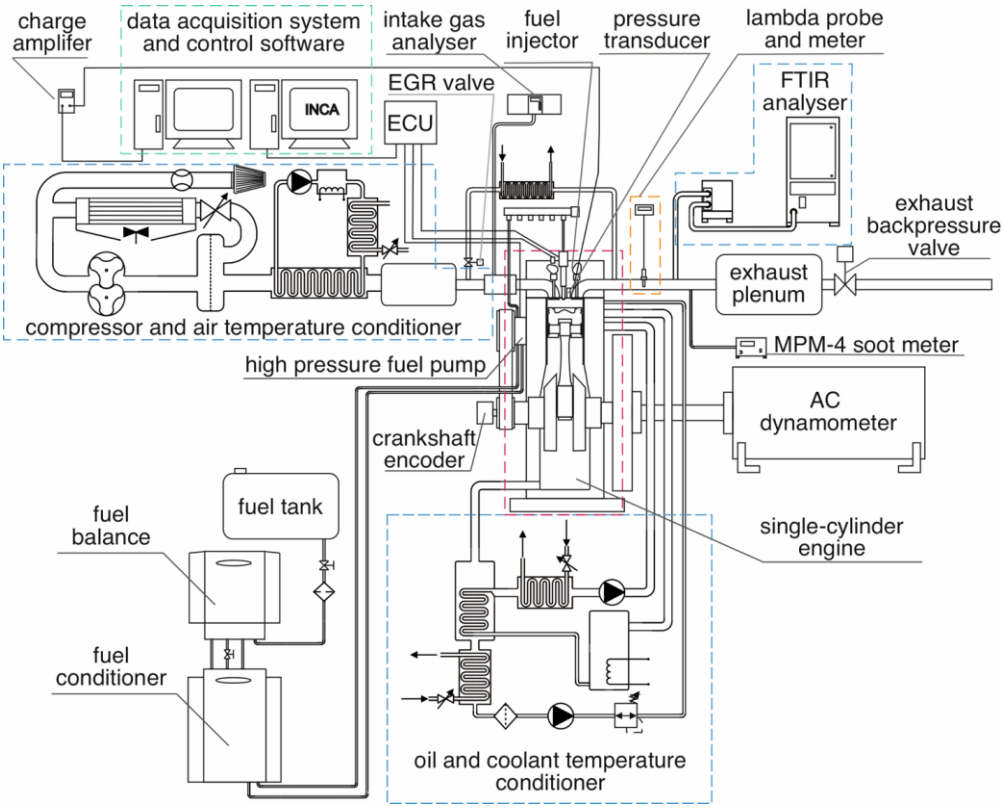


Fig. 1. Diagram of the experimental test stand.

The AVL Sesam FTIR multi-compound gas analysis system was used to determine the composition of exhaust gases (CO, HC, and NOx). A Maha MPM-4 analyser was used to assess exhaust particulate concentrations. The extra air factor was determined using the ETAS LA4 lambda meter and a Bosch LSU 4.2 wideband lambda probe.

2.3. Conditions of the Tests and Engine Outcome Parameters

The tests performed at a CI engine rotational speed of 1500 rpm and the net IMEP for non-EGR operation was around 4.8 bar. The engine run uncharged at ambient pressure in the intake system. At each operating point the engine was thermally stabilized. The lube oil, engine coolant and EGR temperatures were maintained at the same 85 °C. In addition, the fuel temperature was maintained at 30 °C in all experiments and the fuel rail pressure was 80 MPa. In addition, the series of experiments consisted of different strategies and single and split injection as well as EGR sweeps shown in Table 2.

During the experiment, various engine output parameters were measured, including cylinder pressure, heat release rate, mass fraction burnt, and emissions of PM, CO, HC, and NO_X. The cylinder pressure was recorded for 100 consecutive engine cycles at each engine operating point and then averaged. Additionally, pressure was recorded at a fixed crank angle resolution of 0.1 °CA using a piezoelectric transducer installed

directly on the engine.

AVL Boost specialized software was used to accurately process the pressure data. The heat release rate was calculated using the first law of thermodynamics, and the cumulative HRR was used to calculate mass fraction burnt (MFB), which allowed for combustion timing indicators such as the 5, 50, and 95% MFB location (CA5, CA50, and CA95). The start of combustion angle was taken as the 5% MFB, and the combustion time was calculated as the angle between 5% and 95% MFB.

Table 2. Scope of the experiments.

Combustion mode	SOI_1 [°CA]	SOI_2 [°CA]	Injection timing strategy	EGR ratio [%]	EGR ratio strategy	Experiment number (DF)	Experiment number (HVO)
Split injection PPCI	338	356	0	0-42	1-9	1-9	37-45
	342	356	1	0-42	1-9	10-18	46-54
	346	356	2	0-42	1-9	19-27	55-63
Single-pulse CDC	-	354	3	0-42	1-9	28-36	64-72

CO was measured with an accuracy of 0.36%, HC with an accuracy of 0.1-0.49% (depending on type of hydrocarbon species) and NOx with an accuracy of 0.31%. However, the accuracy of PM measurement was 0.1 mg/m^3.

2.4. Type of Fuel

The research was conducted for two test fuels: pure diesel fuel (DF) and and hydrotreated vegetable oil (HVO). Table 3 shows the key physicochemical parameters of the fuels along with the methods of determination and accuracy.

In the case of HVO, that the cetane number is slightly lower DF than HVO. However, the gross and lower heating values are comparable for both fuels.

Table 3. Fuel properties of mixture HVO and DF.

Properties	Device	Method	Accuracy	Fuel	
				HVO	DF
Gross heating value [MJ/kg]	IKA C 5000 calorimeter	DIN 51900-2	130 J/g	47.194	45.894
Lower heating value (LHV) [MJ/kg]				43.737	42.825
Dynamic viscosity at 40 °C [mPa × s]	Anton Paar SVM 3000/G2 Stabinger Viscometer	ASTM D7042	0.1%	2.198	2.412
Kinematic viscosity at 40 °C [mm^2/s]			0.1%	2.876	2.940
Density at 40 °C [g/ml]			0.0002 g/cm^3	0.781	0.820
Flash point [°C]	FP93 5G2 Pensky-Martens analyser	ISO 2719	0.03 °C	66.3	70.5
Cetane number (CN) [-]	PetroSpec analyser	ASTM D613	0.05%	74.5	54.1

2.5. Vibrations and Sound Measurement System

A GRAS 46AE microphone (Frequency range: 3.15 Hz to 20 kHz; Dynamic range: 17 dB (A) to 138 dB; Sensitivity: 50 mV / Pa) was used to measure sound pressure (Fig. 2b (position 1)). Engine vibrations were measured at 3 points (Fig. 2c (positions 2, 3 and 4)) in the vertical (Z), longitudinal (Y) and transverse (X) directions using four Bruel&Kjear 8341 CCLD accelerometers (Frequency range: 0.3 – 10000 Hz; Sensitivity: 10 mV/ms^{-2}) [22]. Data on noise and vibrations were obtained using the Bruel&Kjear Machine Diagnostic Toolbox. The Machine Diagnostics Toolbox consists of Machine Diagnostics Toolbox Type 9727 and the versatile Machine Diagnostics Toolbox Software Bundle Type 7910. Type 9727 includes the multichannel PULSE data acquisition unit Type 3560-B (5-channel) [46].

2.6. Vibration and sound pressure of the engine

Determining vibrations and noise emitted by components of a diesel engine is one of the most difficult environmental tasks, because each engine mechanism affects vibrations and noise separately.

For each experiment, vibration and sound pressure data were collected from the engine unit with a 6.4 kHz sampling frequency for 1 s.

Figure 3 illustrates vibration accelerations and sound pressure using D100 and HVO fuels, respectively, under planned conditions (Table 1). The data received were used in further statistical analysis.

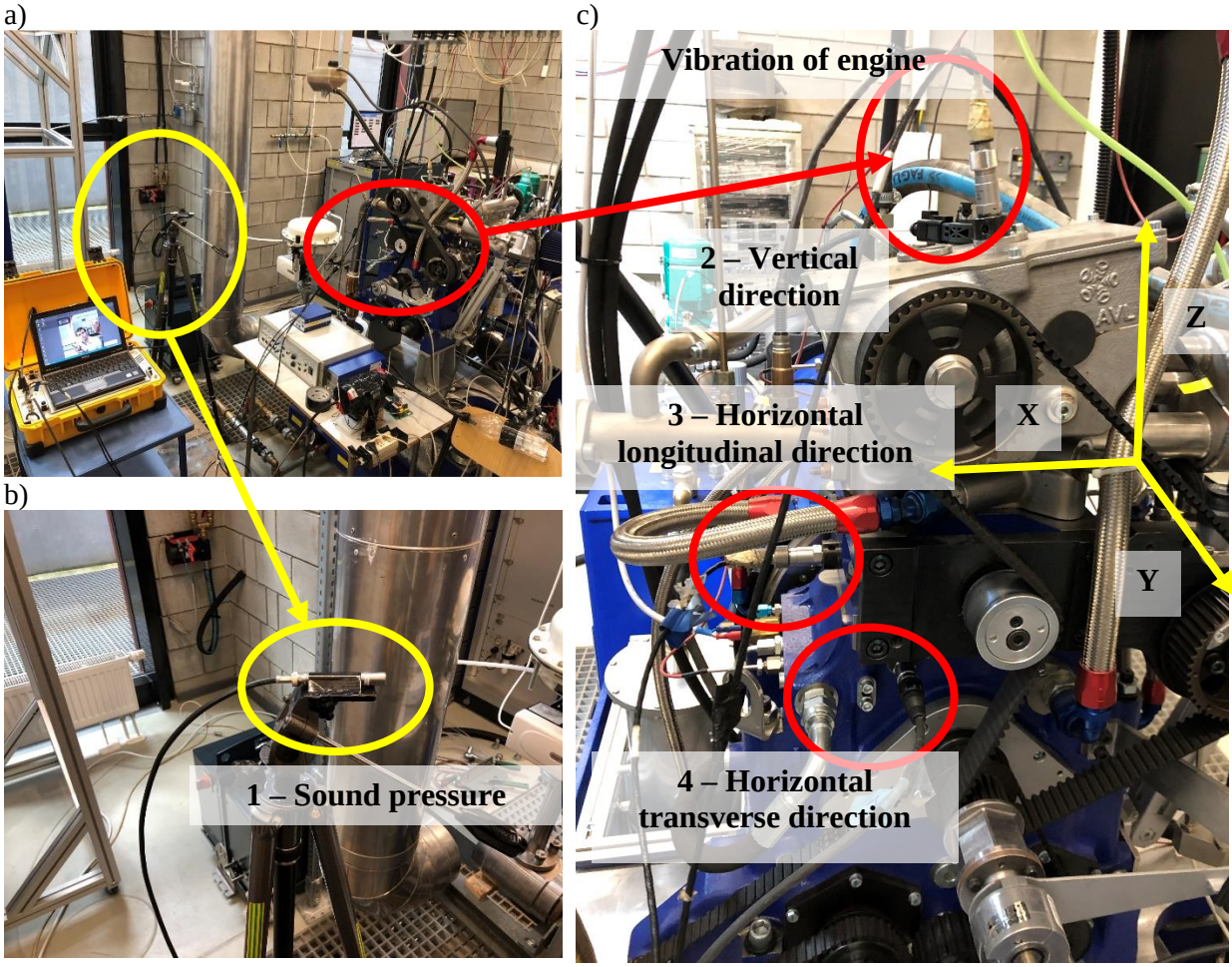


Fig. 2. Points of measurement of vibrations and sound pressure: a) general view of the engine being analyzed and points of measurement of vibrations; b) a microphone to measure sound pressure (point 1); c) accelerometers to measure vibrations (point 2, 3 and 4) in X, Y and Z directions.

2.7. Statistical Analysis

Univariate linear regression model (LRM) was used to identify statistically significant prognostic factors for engine outcome parameters:

$$Y_i^E = \alpha + \beta T_i + \varepsilon,$$

where Y_i^E – engine outcome variable (dependent variable), E – type of engine outcome variable, i – number of experiment, T_i^{RMS} – independent prognostic factor, α – regression intercept value, β – regression parameter for independent prognostic variable, ε – random error.

Independent prognostic factor (main predictors) of LRM were type of fuel, injection timing, EGR ratio, sound pressure (SP), vertical vibration (VV), horizontal vibration at point 1 (HV1) and horizontal vibration at point 2 (HV2). Dependent LRM variables were engine outcome parameters: $\max_{\_press}$, $\max_{\_press\_loc}$, $\mathrm{HRR}_{\max}$, CA05, CA50, CA05_95, CA95, PM.

P, VV, HV1 and HV2 were aggregated with root mean square (RMS) estimate for each experiment:

$$\mathrm{RMS}_j^i = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2},$$

where j – experiment number, i – defines measured parameter (i.e., SP, VV, HV1 and HV2), N – measured parameter sample size, x – value of measured parameter.

In addition, a three-step multivariate LRM algorithm was

investigated to clarify independent prognostic factors for engine outcome parameters. The first step called the full prognostic model was performed to evaluate the prognostic impact of the parameters interactions, the second step called the complete prognostic model was performed to evaluate the prognostic impact of the main predictors after eliminating the interactions of statistically not significant parameters, and the third step called the optimal prognostic model was performed to evaluate the prognostic impact of statistically significant parameters and their significant interactions.

Multivariate LRM were based on analysis of covariance (ANCOVA) model:

$$Y_i^E = \alpha + \beta T_i^{\mathrm{RMS}} + \gamma Z_j + \mu(T_i^{\mathrm{RMS}} * Z_j) + \varepsilon,$$

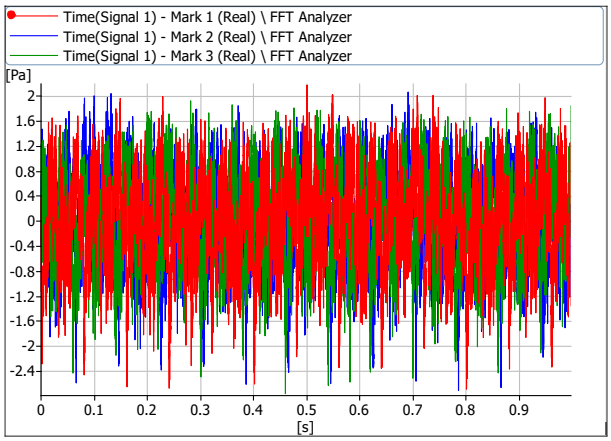
where Y_i^E – engine outcome variable (dependent variable), E – type of engine outcome variable, i – number of experiment, T_i^{RMS} – RMS estimate of sound pressure or vibration data (covariate), Z_j – categorical variable with j levels (independent factor), α – regression intercept value, β – regression parameter for covariate, γ – regression parameter for independent variable, μ - regression parameter for covariate and independent variable interaction, ε – random error.

Accuracy between the optimal prognostic model and real data was evaluated using mean absolute percentage error (MAPE):

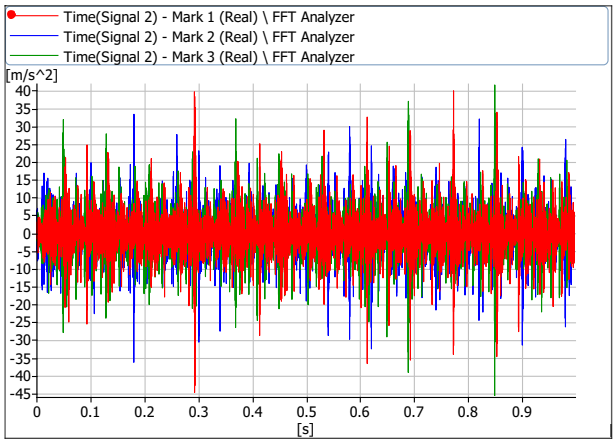
$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{R_i^E - \widehat{Y_i^E}}{R_i^E}\right|,$$

where R_i^E – real engine outcome parameter value for ith experiment, $\widehat{Y_i^E}$ – prognostic engine outcome parameter value for ith experiment, E – engine outcome parameter.

D100

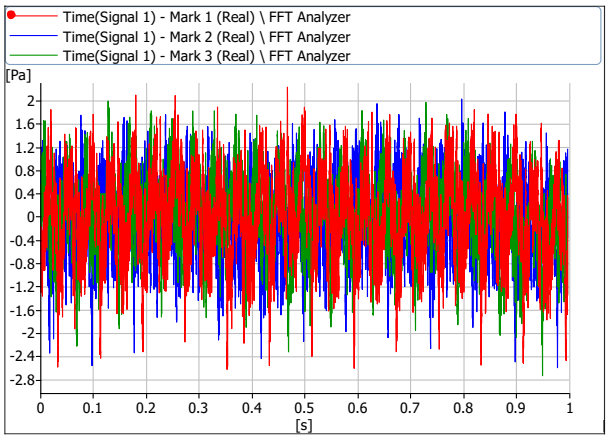


a)

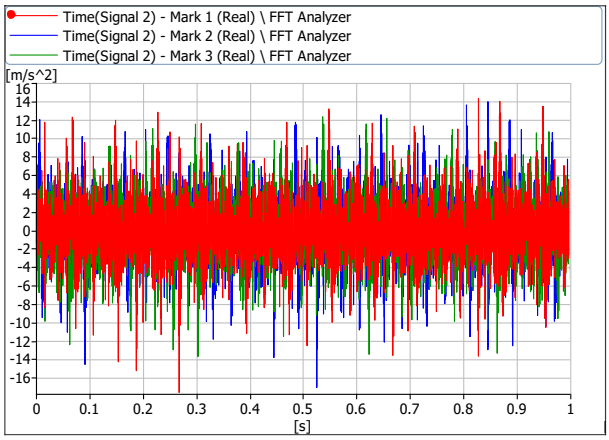


b)

HVO



a1)



b1)

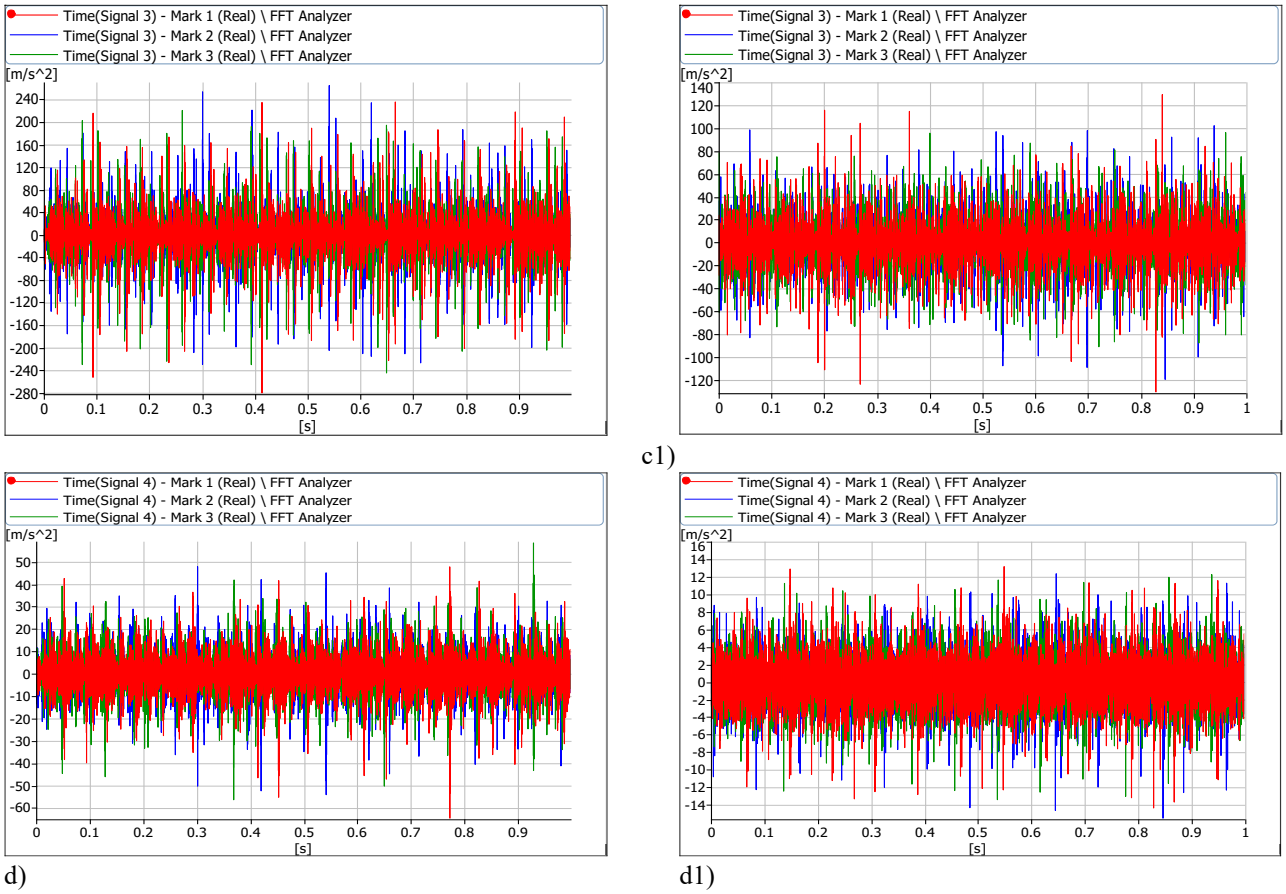


Fig. 3. Typical results of measurement of sound pressure (a and a1) and vibrations (b and b1, c and c1, d and d1) respectively, marking: red - D100 EGR00 and HVO EGR00; blue - D100 EGR03 and HVO50 EGR03; green - D100 EGR07 and HVO50 EGR07.

Before LRM analysis descriptive statistical analysis was performed for all dependent and independent variables. Mean and standard deviation were presented in total sample size and separately in fuel type, injection timing, and EGR ratio subgroups. Statistical differences between subgroups were evaluated using t-test (if $k = 2$) and ANOVA ($k > 2$) statistical tests.

A two-tailed p-value < 0.1 was considered to be statistically significant threshold for independent regression model parameters.

3. Results and Discussion

3.1. Descriptive Statistics of Exhausted Emission Parameters

Descriptive analysis performed for engine outcome parameters, sound pressure and vibration data in type of fuel, injection timing and EGR ratio subgroups.

The mean values of engine outcome parameters and sound pressure RMS were not different in DF and HVO fuel subgroups (Table 4). Statistically significantly higher mean estimates were observed in DF for VV, HV1, and HV2. Average of VV and HV2 RMS were 1.7 and 2.8 times higher ($p < 0.001$) in DF group, while HV1 RMS was only 1.3 times higher in DF group, but the difference remained statistically significant.

Most engine outcome parameters, sound pressure and vibration data were distributed differently in injection timing subgroups (Table 5). The mean values of $max_{\_press}$, $max_{\_press\_loc}$, CA50, CO, HC and NO_X were though all injection timing subgroups. The major impact for mean values had 3rd subgroup of injection timing. The average of HRR_{max} and vibration data were doubled in 3rd injection timing subgroup compared with the rest three subgroups (0, 1 and 2). On the contrary, CA05-95 was three times lower in 3rd injection timing subgroup then in others. Interestingly, average PM estimates were non-stable going though all injection timing subgroups, i.e., mean (SD) of PM in 0, 1, 2, and 3 injection timing subgroups were 8.7 (2.85), 7.8 (2.45), 25.9 (7.97) and 14.9 (30.32) ($p = 0.003$), respectively.

Table 4. Descriptive statistics of engine outcome parameters and vibration data in fuel subgroups.

Parameter	Total Mean (SD)	DF Mean (SD)	HVO Mean (SD)	P value
Max pressure, Mpa	5.6 (0.73)	5.6 (0.73)	5.7 (0.73)	0.587
Max press location, °CA	369.3 (3.09)	369.3 (3.20)	369.2 (3.02)	0.847
HRR max, J/°CA	70.7 (31.47)	69.0 (30.89)	72.4 (32.40)	0.650
CA05, °CA	362.1 (4.66)	362.2 (4.90)	362.0 (4.48)	0.857
CA50, °CA	368.2 (4.47)	368.5 (5.10)	368.0 (3.80)	0.591
CA05-95, °CA	25.8 (11.84)	26.7 (12.23)	25.0 (11.55)	0.549
CA95, °CA	388.0 (11.42)	388.9 (12.22)	387.0 (10.65)	0.487
PM, mg/m^3	14.3 (17.07)	14.8 (17.01)	13.8 (17.35)	0.810
CO, ppm	943.3 (936.54)	981.4 (1010.03)	905.2 (869.57)	0.732
HC, ppm	391.3 (418.16)	423.9 (545.13)	358.8 (235.31)	0.514
NO_X, ppm	329.0 (278.61)	333.9 (282.91)	324.1 (278.17)	0.882
Sound pressure RMS	0.8 (0.09)	0.8 (0.09)	0.8 (0.10)	0.945
Vertical vibration RMS	4.9 (2.72)	6.2 (3.07)	3.6 (1.42)	<0.001
Horizontal vibration 1 RMS	26.6 (9.70)	29.6 (7.51)	23.6 (10.74)	0.007
Horizontal vibration 2 RMS	7.4 (6.82)	10.9 (8.08)	3.9 (2.06)	<0.001

Table 5. Descriptive statistics of engine outcome parameters and vibration data in injection timing subgroups.

Parameter	0 Mean (SD)	1 Mean (SD)	2 Mean (SD)	3 Mean (SD)	P value
Max pressure [Mpa]	5.6 (0.54)	5.6 (0.72)	5.6 (0.42)	5.8 (1.09)	0.717
Max press location [°CA]	369.8 (2.62)	369.2 (2.08)	369.4 (2.28)	368.6 (4.76)	0.728
HRRmax [J/°CA]	54.6 (11.96)	69.1 (12.85)	51.8 (11.48)	107.5 (40.15)	<0.001
CA05 [°CA]	360.4 (4.19)	360.9 (4.51)	360.6 (4.03)	366.7 (2.60)	<0.001
CA50 [°CA]	367.5 (2.81)	367.9 (4.07)	367.3 (2.62)	370.2 (6.80)	0.169
CA05-95 [°CA]	29.4 (1.09)	29.1 (2.47)	34.7 (1.53)	10.1 (14.37)	<0.001
CA95 [°CA]	389.8 (4.01)	389.9 (4.89)	395.3 (5.05)	376.9 (16.86)	<0.001
PM [mg/m^3]	8.7 (2.85)	7.8 (2.45)	25.9 (7.97)	14.9 (30.32)	0.003
CO [ppm]	772.2 (745.77)	978.0 (859.24)	787.3 (888.28)	1235.7 (1196.33)	0.418
HC [ppm]	324.3 (177.68)	357.3 (238.18)	295.7 (150.31)	588.1 (750.16)	0.137
NO_X [ppm]	288.6 (217.33)	296.0 (225.71)	265.2 (210.95)	466.3 (392.90)	0.112
Sound pressure RMS	0.7 (0.02)	0.7 (0.03)	0.7 (0.03)	0.9 (0.10)	<0.001
Vertical vibration RMS	3.7 (0.89)	4.6 (1.67)	3.6 (0.90)	7.6 (3.91)	<0.001
Horizontal vibration 1 RMS	21.7 (4.47)	24.2 (4.93)	21.9 (4.69)	38.6 (10.99)	<0.001
Horizontal vibration 2 RMS	4.9 (1.99)	6.3 (3.28)	4.9 (2.30)	13.6 (10.92)	<0.001

Descriptive analysis in EGR ratio subgroups revealed contrary tendency for parameter distributions than it was observed injection timing subgroups (Table 6). $max_{\_press}$, $max_{\_press\_loc}$, CA05, CA50, CA95, CO, HC and NO_X parameters had statistically different distributions in EGR ratio subgroups. Only CA05 mean estimate was statistically significantly different in between injection and EGR ratio subgroups. The major impact for higher mean values of CA50 and PM was observed in 42 EGR ratio subgroup. Mean of PM was almost 3 times higher in 42 EGR ratio subgroup but statistical significance was not reached due to high level of scattering of raw values. The same findings with high level scattering of raw values were observed in HRR_{max} and CA05-95.

3.2. Univariate regression analysis

Univariate LRM analysis were performed for engine outcome parameters as dependent parameter and type of fuel, injection timing, EGR ratio and RMS of vibration and sound pressure

data as independent factors (Table 7). Type of fuel had no statistically significant impact for all dependent parameters with R^2 lower than 1 and p-value > 0.1.

Table 6. Descriptive statistics of engine outcome parameters and vibration data in EGR ratio subgroups.

Parameter	0 Mean (SD)	3 Mean (SD)	7 Mean (SD)	10 Mean (SD)	14 Mean (SD)	21 Mean (SD)	28 Mean (SD)	35 Mean (SD)	42 Mean (SD)	P value
Max pressure [Mpa]	6.1 (0.38)	6.1 (0.36)	6.1 (0.34)	6.0 (0.31)	6.0 (0.27)	5.8 (0.16)	5.6 (0.20)	5.1 (0.38)	4.0 (0.43)	<0.001
Max press location [°CA]	368.0 (0.56)	368.1 (0.39)	368.2 (0.27)	368.3 (0.32)	368.4 (0.48)	368.9 (0.96)	370.4 (1.75)	373.0 (2.32)	370.3 (7.81)	0.010
HRRmax [J/°CA]	70.9 (43.73)	72.0 (43.26)	72.9 (40.82)	72.9 (34.43)	77.2 (35.66)	80.1 (24.21)	79.3 (11.00)	69.7 (5.27)	41.7 (13.94)	0.389
CA05 [°CA]	358.7 (3.59)	358.7 (3.85)	359.1 (3.66)	359.7 (3.55)	360.5 (3.22)	362.1 (2.92)	364.5 (1.97)	366.5 (1.5)	369.4 (2.23)	<0.001
CA50 [°CA]	365.5 (0.23)	365.7 (0.29)	365.8 (0.39)	366.0 (0.56)	366.3 (0.66)	367.1 (1.04)	368.5 (1.47)	371.2 (2.00)	378.1 (6.19)	<0.001
CA05-95 [°CA]	22.8 (12.87)	24.0 (12.60)	24.1 (13.03)	24.6 (13.03)	24.6 (13.32)	25.5 (13.20)	25.2 (11.11)	26.3 (7.61)	35.5 (9.75)	0.613
CA95 [°CA]	381.6 (9.45)	382.6 (8.91)	383.2 (9.51)	384.4 (9.64)	385.1 (10.20)	387.5 (10.57)	389.7 (9.44)	392.8 (6.46)	405.0 (11.56)	<0.001
PM [mg/m^3]	10.6 (5.29)	10.6 (5.75)	11.9 (6.85)	13.0 (8.09)	14.2 (11.06)	14.8 (13.88)	12.7 (13.75)	11.2 (12.81)	30.0 (42.19)	0.443
CO [ppm]	287.2 (66.10)	318.8 (85.15)	356.1 (100.75)	408.8 (134.54)	484.2 (163.70)	715.5 (226.46)	1090.3 (286.15)	1713.8 (315.44)	3114.8 (739.07)	<0.001
HC [ppm]	210.1 (7.34)	220.1 (14.48)	228.8 (19.12)	242.0 (23.89)	261.5 (38.36)	295.1 (71.40)	364.5 (108.97)	520.9 (167.11)	1179.0 (916.76)	<0.001
NO_X [ppm]	759.1 (223.55)	622.9 (170.57)	520.0 (151.03)	420.5 (114.33)	320.2 (92.01)	172.9 (35.00)	89.1 (14.72)	41.1 (10.85)	15.3 (4.29)	<0.001
Sound pressure RMS	0.8 (0.14)	0.8 (0.14)	0.8 (0.12)	0.8 (0.11)	0.8 (0.11)	0.8 (0.06)	0.8 (0.02)	0.7 (0.02)	0.7 (0.05)	0.852
Vertical vibration RMS	5.7 (3.79)	5.5 (3.67)	5.7 (3.94)	5.1 (2.84)	5.2 (2.68)	5.3 (2.60)	4.5 (1.26)	4.0 (1.00)	3.2 (0.65)	0.648
Horizontal vibration 1 RMS	30.1 (13.22)	29.3 (12.60)	29.2 (12.12)	28.6 (11.19)	28.3 (10.38)	26.3 (7.69)	24.4 (4.56)	22.3 (4.55)	20.9 (6.12)	0.489
Horizontal vibration 2 RMS	9.4 (9.89)	9.0 (9.92)	8.8 (9.51)	8.3 (8.19)	8.1 (6.89)	7.3 (4.59)	6.3 (3.11)	5.3 (2.63)	4.2 (1.89)	0.837

Table 7. Univariate regression model for engine outcome parameters.

Independent factor	Dependent parameter										
	Max press	Max press loc	HRR max	CA05	CA50	CA05-95	CA95	PM	CO	HC	NOx
	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value	R^2, p-value
Fuel	<1, 0.587	<1, 0.847	<1, 0.650	<1, 0.857	<1, 0.591	<1, 0.549	1, 0.487	<1, 0.810	<1, 0.732	<1, 0.513	<1, 0.882
Injection timing	2, 0.717	2, 0.728	50, <0.001	33, <0.001	7, 0.169	63, <0.001	36, <0.001	18, 0.003	<1, 0.418	8, 0.137	8, 0.112
EGR ratio	82, <0.001	26, 0.010	12, 0.389	62, <0.001	77, <0.001	9, 0.613	37, <0.001	11, 0.463	90, <0.001	50, <0.001	84, <0.001
Sound pressure RMS	23, <0.001	3, 0.137	79, <0.001	13, 0.002	1, 0.502	67, <0.001	49, <0.001	3, 0.133	1, 0.357	<1, 0.690	31, <0.001
Vertical vibration RMS	22, <0.001	1, 0.374	54, <0.001	4, 0.090	3, 0.133	44, <0.001	37, <0.001	6, 0.043	4, 0.108	2, 0.194	27, <0.001
Horizontal vibration 1 RMS	27, <0.001	2, 0.212	68, <0.001	5, 0.050	3, 0.153	59, <0.001	49, <0.001	6, 0.046	4, 0.097	1, 0.328	36, <0.001
Horizontal vibration 2 RMS	16, <0.001	1, 0.434	41, <0.001	4, 0.092	2, 0.252	34, <0.001	27, <0.001	4, 0.114	2, 0.200	1, 0.314	23, <0.001

Six independent prognostic factors (i.e., all independent prognostic factors except type of fuel) had prognostic impact for CA05 and CA95 in univariate LRM. Five statistically significant independent prognostic factors had max_press, HRRmax, CA05-95, and NO_X. However, max_press_loc, CA50, and HC had only one prognostic factor EGR ratio with $R^2 = 26$ ($p = 0.010$), $R^2 = 77$ ($p < 0.001$), and $R^2 = 50$ ($p < 0.001$), respectively. There were no prognostic factors which had statistically significant impact for all of engine parameters.

3.3. Multivariate regression model building

Optimal multivariate LMR should contain minimal number of independent factors due to reduction of model errors. For this reason, multivariate model should contain only these independent prognostics have the major impact for dependent variable. Multivariate LRM building is stepwise statistical procedure with following steps: (1) evaluation of prognostic impact of parameters interaction (Full prognostic model), (2) evaluation of prognostic impact of parameters (main predictors) after elimination of statistically non-significant parameters' interactions (Complete prognostic model) and (3) evaluation of prognostic impact of parameters after elimination of statistically non-significant parameters from complete prognostic model (Optimal prognostic model).

Type of fuel was not included in multivariate LRMs because it showed no evidence to have any impact for all engine outcome parameters. Multivariate LRMs contained interactions composed from sound pressure and vibration data interactions with injection timing and EGR ratio subgroups.

Full prognostic model

The main goal of full regression model is to evaluate prognostic impact for engine outcome parameters. Only 6 interactions between injection timing and vibration parameters were observed to be statistically significant (Table 8). Much more statistically significant interactions identified between EGR ratio and vibration parameters. Four and three interactions between EGR ratio and vibration parameters were statistically significant for Max_press, HRRmax, and NO_X prognostic models, respectively.

In summary, engine outcome parameters CA05, CA50, CA95, PM, CO, and HC prognostic models had no influence from any interaction parameters, $max_{\_press\_loc}$, CA05-95 models had only one statistically significant interaction parameter and $max_{\_press}$ with HRR_{max} and NO_X models had high impact from interacting parameters (Table 8).

Complete prognostic model

In the second step called complete prognostic model there were eliminated statistically non-significant (i.e., without prognostic impact) interactions from the full regression model (Table 9). Complete prognostic models revealed that high number of statistically non-significant interaction parameters increased model errors and main predictors showed very mild impact for engine parameters. High prognostic influence increment in all engine parameter models observed for injection timing after elimination of statistically non-significant interactions. EGR ratio and sound pressure data also showed high prognostic increment for most of prognostic models. Only vertical and horizontal vibration data remained statistically significant in $max_{\_press}$ and HHR_{max} prognostic models (Table 9). Almost all (except NO_X) interaction parameters with prognostic impact from the previous step remained with slightly higher prognostic impact in complete prognostic model.

Optimal prognostic model

The last third step was optimal prognostic models building and estimation of models' parameters. Optimal prognostic models were made of statistically significant independent factors from complete prognostic models. None of the parameters became without prognostic impact after elimination of statistically non-significant parameters from the complete prognostic models.

Optimal models for $max_{\_press}$, HRR_{max} and CA05 showed very high goodness-of-fit level with $R^2 > 90\%$. Also, high level of model goodness-of-fit was estimated for $max_{\_press\_loc}$, CA50, CA05-95 and CA95 with R^2 between 70%-90%.

Table 8. Multivariate regression model for engine outcome parameters: full prognostic model.

	Max press p-value	**Max press loc** p-value	**HRR max** p-value	**CA05** p-value	**CA50** p-value	**CA05-95** p-value	**CA95** p-value	**PM** p-value	**CO** p-value	**HC** p-value	**NOx** p-value
Model R^2	100	96	100	100	98	98	97	91	99	89	100
Injection timing	0.179	0.237	0.991	0.395	0.772	0.753	0.734	0.581	0.691	0.656	0.130
EGR ratio	0.170	0.160	0.798	0.796	0.992	0.923	0.891	0.552	0.993	0.988	0.121
Sound pressure RMS	0.673	0.237	0.857	0.850	0.812	0.659	0.680	0.434	0.994	0.949	0.973
Vertical vibration RMS	0.003	0.010	0.064	0.092	0.141	0.185	0.134	0.102	0.902	0.888	0.839
Horizontal vibration 1 RMS	0.005	0.307	0.049	0.619	0.959	0.560	0.606	0.526	0.457	0.712	0.725
Horizontal vibration 2 RMS	<0.001	0.035	0.010	0.072	0.141	0.386	0.283	0.196	0.509	0.954	0.922
Injection timing interaction with											
Sound pressure RMS	0.472	0.245	0.904	0.467	0.702	0.677	0.611	0.521	0.689	0.596	0.026
Vertical vibration RMS	0.252	0.650	0.778	0.417	0.643	0.087	0.114	0.623	0.299	0.227	0.058
Horizontal vibration 1 RMS	0.016	0.253	0.049	0.673	0.678	0.841	0.790	0.563	0.886	0.996	0.135
Horizontal vibration 2 RMS	0.626	0.348	0.612	0.573	0.608	0.129	0.158	0.356	0.394	0.359	0.098
EGR ratio interaction with											
Sound pressure RMS	0.046	0.042	0.026	0.246	0.322	0.756	0.629	0.488	0.756	0.662	0.039
Vertical vibration RMS	0.007	0.119	0.161	0.609	0.833	0.672	0.611	0.387	0.963	0.998	0.097
Horizontal vibration 1 RMS	<0.001	0.791	0.045	0.837	0.978	0.963	0.960	0.974	0.981	0.991	0.012
Horizontal vibration 2 RMS	<0.001	0.163	0.005	0.421	0.403	0.671	0.569	0.314	0.961	0.970	0.311

Table 9. Multivariate regression model for engine outcome parameters: complete prognostic model.

	Max press p-value	**Max press loc** p-value	**HRR max** p-value	**CA05** p-value	**CA50** p-value	**CA05-95** p-value	**CA95** p-value	**PM** p-value	**CO** p-value	**HC** p-value	**NOx** p-value
Model R^2	100	71	99	97	98	87	77	37	96	64	100
Injection timing	<0.001	0.079	<0.001	<0.001	<0.001	0.002	0.008	0.002	<0.001	0.001	0.003
EGR ratio	0.811	<0.001	0.001	<0.001	<0.001	0.021	<0.001	0.719	<0.001	<.001	0.762
Sound pressure RMS	0.890	0.096	0.181	0.001	0.017	0.034	0.097	0.525	0.007	0.118	0.082
Vertical vibration RMS	0.003	0.357	0.529	0.987	0.186	0.343	0.508	0.618	0.479	0.279	0.793
Horizontal vibration 1 RMS	<0.001	0.920	<0.001	0.656	0.759	0.314	0.639	0.393	0.965	0.972	0.557
Horizontal vibration 2 RMS	<0.001	0.476	0.022	0.589	0.165	0.304	0.446	0.605	0.403	0.325	0.801
Injection timing interaction with											
Sound pressure RMS	-	-	-	-	-	-	-	-	-	-	0.002
Vertical vibration RMS	-	-	-	-	-	<0.001	-	-	-	-	0.276
Horizontal vibration 1 RMS	<0.001	-	<0.001	-	-	-	-	-	-	-	-
Horizontal vibration 2 RMS	-	-	-	-	-	-	-	-	-	-	0.378
EGR ratio interaction with											
Sound pressure RMS	0.001	<.0001	<0.001	-	-	-	-	-	-	-	0.376
Vertical vibration RMS	0.011	-	-	-	-	-	-	-	-	-	0.448
Horizontal vibration 1 RMS	<0.001	-	0.006	-	-	-	-	-	-	-	0.770
Horizontal vibration 2 RMS	<0.001	-	0.001	-	-	-	-	-	-	-	-

Developed models showed high level performance and accuracy. The least MAPE values (<1%) was observed in $max_{\_pressure}$, $max_{\_pressure\_loc}$, CA05, CA50, and CA95 (Fig. 4). In model realization graphs it can be seen that these 5 dependent parameters real and optimal (prognostic) value have very high coincidence. Moderate MAPE were evaluated in HRR_{max}, CA05-95, CO, and HC, the values were 5.1%, 17.8%, 13.1%, 13.1%, and 28.4%, respectively. The worst MAPE (>100%) was observed in PM and NO_X (Fig. 4).

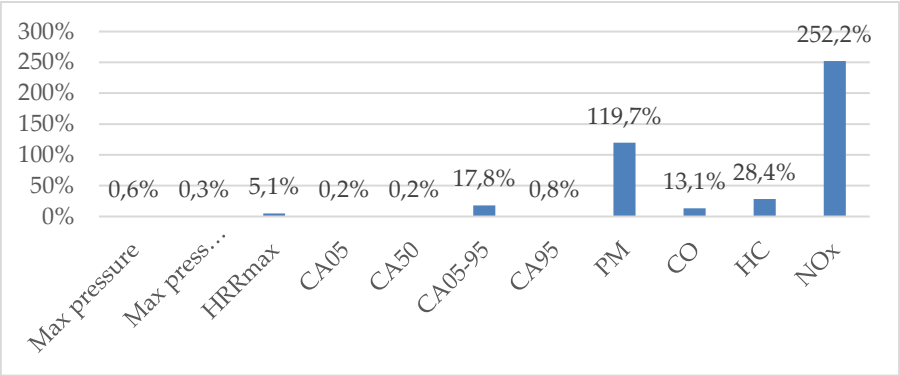
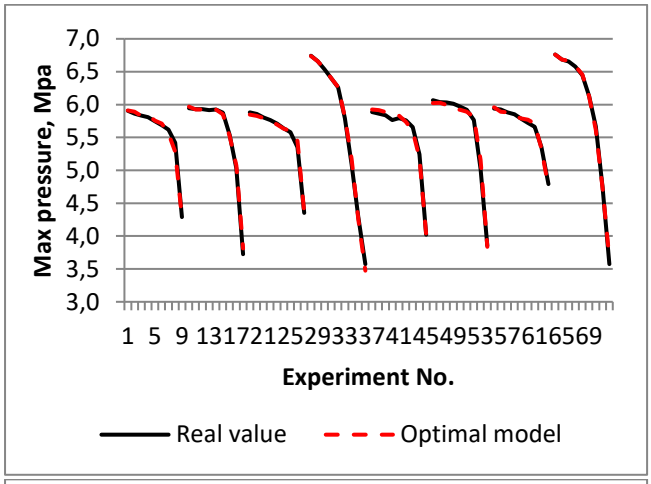
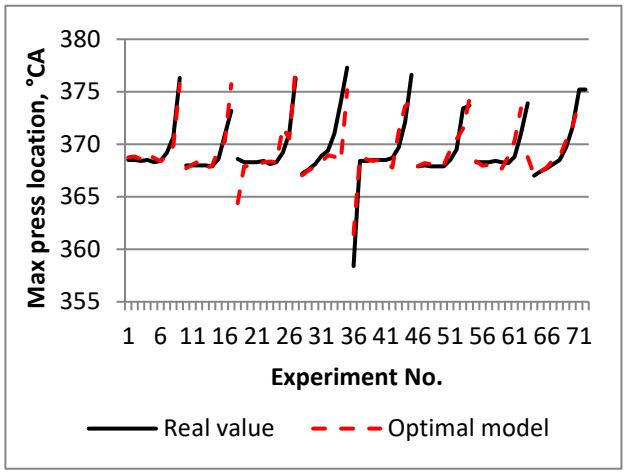
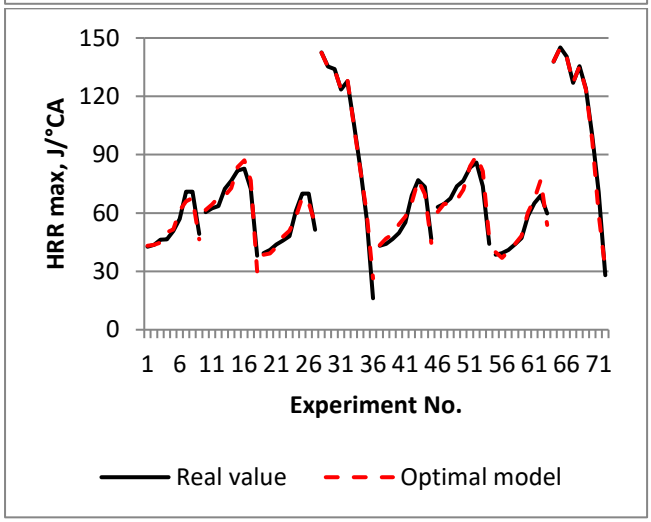


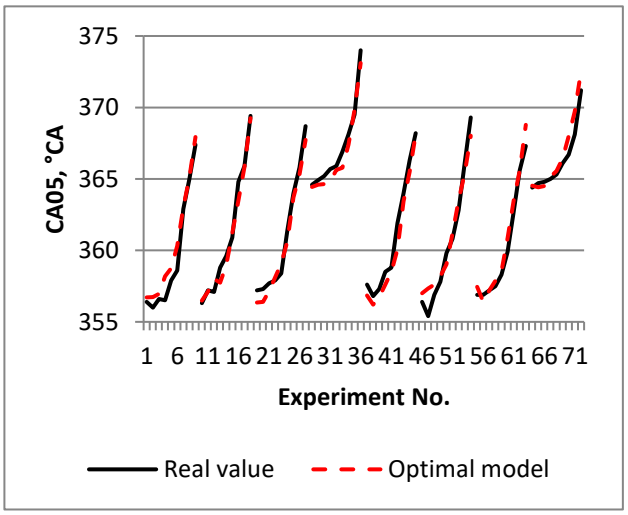
Fig. 4. Mean absolute percentage error for engine outcome parameters calculated from ANCOVA model.

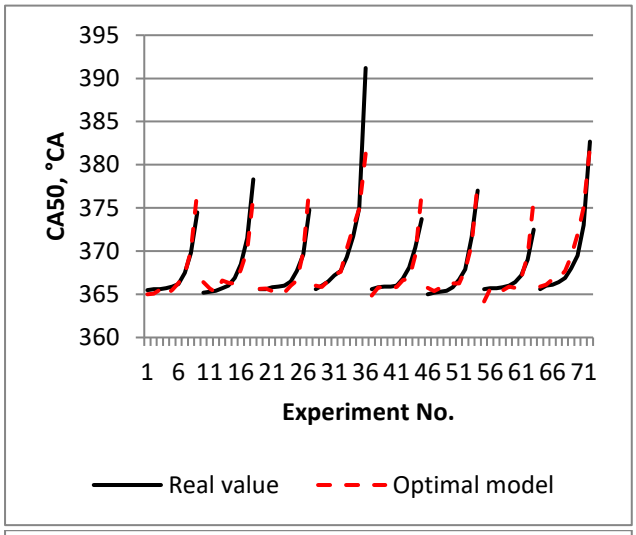
According the literature the XGBoost-based PIDM technique has higher performance in terms of computational efficiency and accuracy, hence establishing its dominance in this domain [24]. Realization of optimal prognostic model and real engine parameter values shown in Fig. 5.

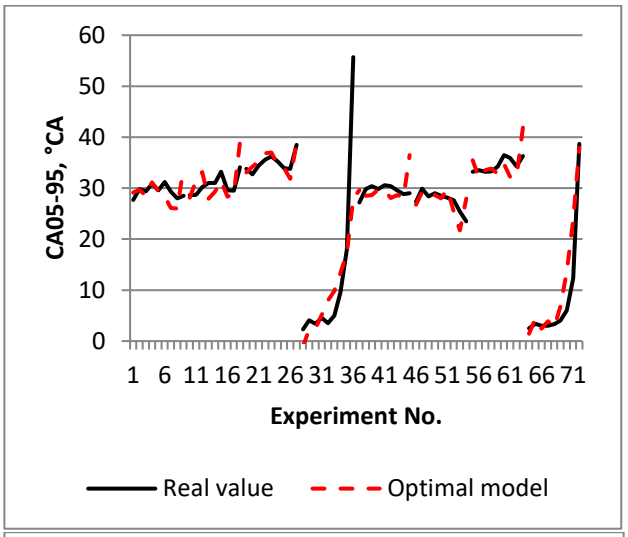


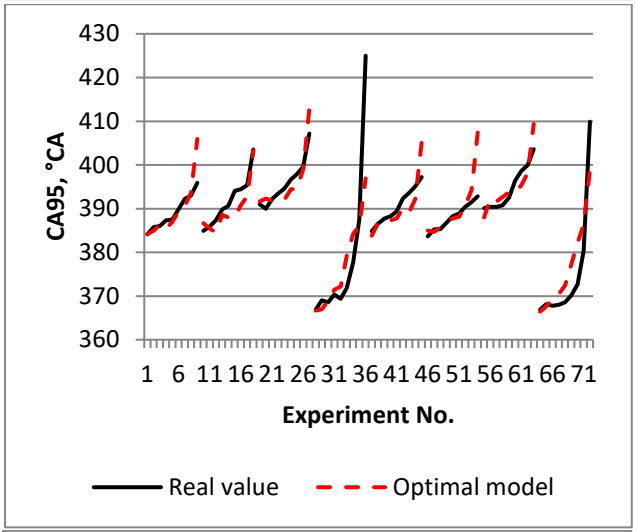


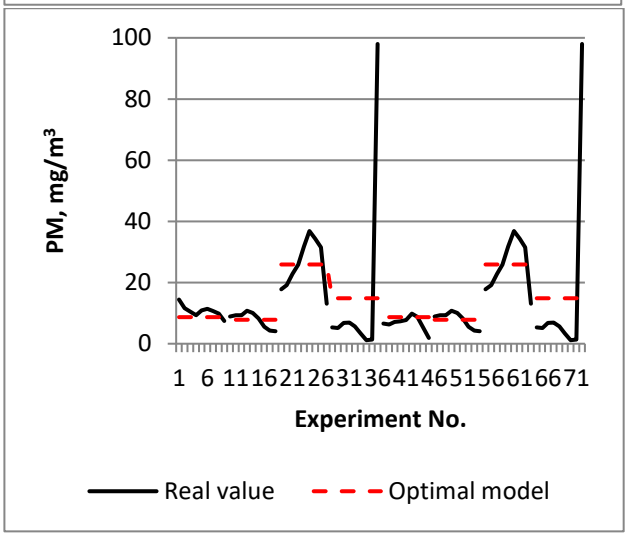


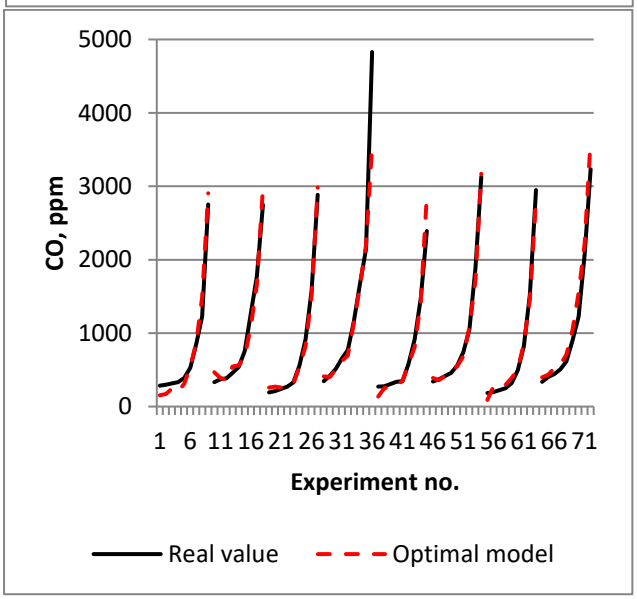


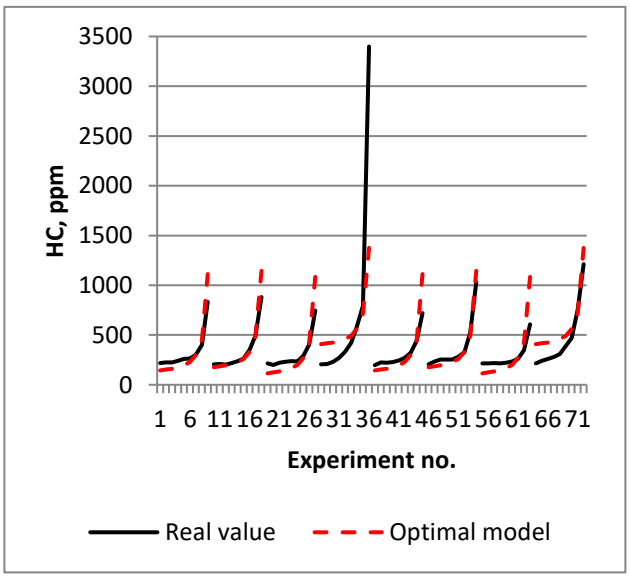

CA50, °CA
395
390
385
380
375
370
365
360
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment No.
Real value
Optimal model


CA05-95, °CA
60
50
40
30
20
10
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment No.
Real value
Optimal model


CA95, °CA
430
420
410
400
390
380
370
360
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment No.
Real value
Optimal model


PM, mg/m³
100
80
60
40
20
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment No.
Real value
Optimal model


CO, ppm
5000
4000
3000
2000
1000
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment no.
Real value
Optimal model


HC, ppm
3500
3000
2500
2000
1500
1000
500
0
1 6 11 16 21 26 31 36 41 46 51 56 61 66 71
Experiment no.
Real value
Optimal model

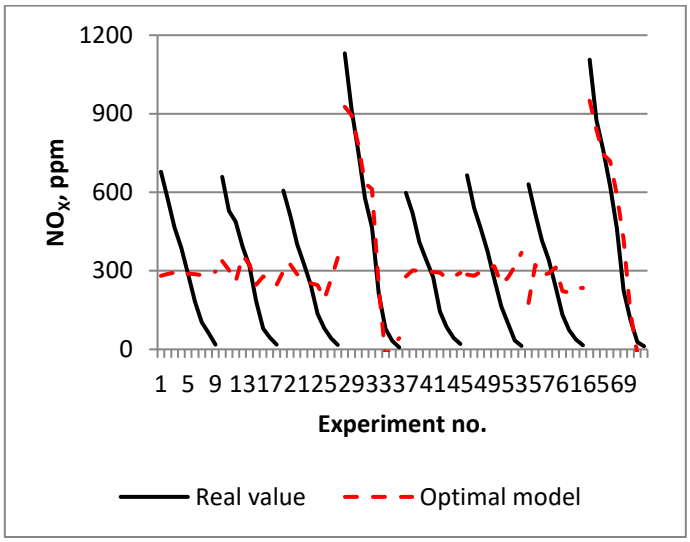


Fig. 5. Engine outcome parameters real and prognostic values: a) Max pressure, MPa; b) Max press location, °CA; c) HRR max, J/°CA; d) CA05, °CA; e) CA50, °CA; f) CA05-95, °CA; g) CA95, °CA; h) PM, mg/m^3; i) CO, ppm; j) HC, ppm; k) NO_X, ppm..

The given Experiment number corresponds to the data presented in Table 1. It can be seen that by drawing the prognostic model and real values neither the injection timing nor the EGR ratio had any major negative effect on the prediction of $max_{\_press}$, $max_{\_press\_loc}$, HRR_{max}, CA05, CA50, CA05-95, CA95, CO, and HC.

PM prognostic values could be treated as acceptable, however prognostic model was not able to correctly predict PM at highest injection and EGR ratio (Fig. 5h). NO_X remained unpredictable using optimal prognostic model. The study suggests that AI-based virtual sensors can improve NOx monitoring precision in diesel engines, potentially reducing emissions and improving air quality [14].

4. Conclusion

The reliability of engine assessment was related to the selection of optimal prediction model for exhausted emission parameters and to identify the best predictable exhausted emission parameters for single-cylinder four-stroke engine, respectively. Statistical analysis of experimental data revealed that optimal prediction model was able to predict the majority of engine emissions with high level of prediction accuracy: $max_{\_press}$, $max_{\_press\_loc}$, HRR_{max}, CA05, CA50, CA05-95, CA95, CO, and HC. Prediction accuracy level was extremely high with MAPE < 1% for $max_{\_press}$, $max_{\_press\_loc}$, CA05, CA50, and CA95. NO_X remained unpredictable using optimal prediction model. Interestingly, in the full prediction model 6 out of 8 interactions were statistically significant in NO_X predictive model. After elimination of non-significant interactions in NO_X complete prediction model only interaction between injection timing and sound pressure has left as statistically relevant. These unstable results implies that such prediction model building cannot be applied for NO_X and needs to investigate new prediction model building techniques.

Subgroup analysis showed that engine emissions between DF and HVO50 remained statistically not different. Higher trend of mean (DF vs. HVO50) values for PM (14.8 vs. 13.8), CO (981.4 vs. 905.2), HC (423.9 vs. 358.8), and NO_X (333.9 vs. 324.1) were observed in DF, but statistical significance was not reached. On the contrary, the mean of VV, HV1, and HV2 were extremely higher in DF with strong statistical significance.

Injection timing had more influence for engine emissions parameters distribution. HRRmax, CA05, CA05-95, CA95, and PM had statistically different means in different injection timing subgroups ($p < 0.1$). There was also demonstrated that emissions' mean values were not a monotonic function of injection timing. Sound pressure and vibration data had statistically different RMS mean values in injection timing. The largest mean values were observed in 3° injection timing and it was 1.3, 2.1, 1.8, and 2.8 times higher than at 0° for sound pressure, VV, HV1, and HV2.

EGR ratio had major influence for engine emissions parameters distribution. Only means of HRR_{max}, CA05-95, and PM remained statistically not different through all EGR ratio subgroups. Increased EGR ratio was related to the decreased NO_X values while other emission parameters increased together

with higher EGR ratio. Neither sound pressure nor vibration data were related with the EGR ratio subgroup ($p > 0.1$).

The study found that incorporating sound pressure and vibration data in predictive models improves emission prediction accuracy for biodiesel fuels. Optimal linear regression models can accurately predict emissions by considering fixed engine parameters. The results provide a reliable and cost-effective method to evaluate alternative fuel mixtures on diesel engine parameters.

Acknowledgments

This paper has received funding under postdoctoral fellowship project from the Research Council of Lithuania (LMTLT), agreement No [S-PD-22-81]. The authors thank the AVL company for the opportunity to use the engine simulation tool AVL BOOST, which was used to analyse the combustion process and present the results. A cooperation agreement has been concluded between the Faculty of the Transport Engineering of Vilnius Tech University and AVL Advanced Simulation Technologies.

References

1. Aatola H, Larmi M, Sarjovaara T, Mikkonen S. Hydrotreated Vegetable Oil (HVO) as a Renewable Diesel Fuel: Trade-off between NOx, Particulate Emission, and Fuel Consumption of a Heavy Duty Engine. 2008. doi:https://doi.org/10.4271/2008-01-2500, https://doi.org/10.4271/2008-01-2500.
2. Alonso J S J, Sastre J A L, Romero-Ávila C, López E. A note on the combustion of blends of diesel and soya, sunflower and rapeseed vegetable oils in a light boiler. Biomass and Bioenergy 2008; 32(9): 880–886, https://doi.org/10.1016/j.biombioe.2008.01.007.
3. Alruqi M, Sharma P, Deepanraj B, Shaik F. Renewable energy approach towards powering the CI engine with ternary blends of algal biodiesel-diesel-diethyl ether: Bayesian optimized Gaussian process regression for modeling-optimization. Fuel 2023; 334: 126827, https://doi.org/10.1016/j.fuel.2022.126827.
4. Arthanarisamy M, Alagumalai A, Natarajan N. Biodiesel as an Alternative Transportation Fuel in Diesel Engines An in-Depth Study on Biodiesel Performance. 1. Auflage, neue Ausgabe. Saarbrücken, LAP LAMBERT Academic Publishing: 2016.
5. Banković-Ilić I B, Stojković I J, Stamenković O S et al. Waste animal fats as feedstocks for biodiesel production. Renewable and sustainable energy reviews 2014; 32: 238–254. https://doi.org/10.1016/j.rser.2014.01.038
6. Bocchetti D, Giorgio M, Guida M, Pulcini G. A competing risk model for the reliability of cylinder liners in marine Diesel engines. Reliability Engineering & System Safety 2009; 94(8): 1299–1307, https://doi.org/10.1016/j.ress.2009.01.010.
7. Borucka A, Kozłowski E, Antosz K, Parczewski R. A New Approach to Production Process Capability Assessment for Non-Normal Data. Applied Sciences 2023; 13(11): 6721, https://doi.org/10.3390/app13116721.
8. Borucka A, Kozłowski E, Parczewski R et al. Supply Sequence Modelling Using Hidden Markov Models. Applied Sciences 2022; 13(1): 231, https://doi.org/10.3390/app13010231.
9. Borucka A, Wiśniowski P, Mazurkiewicz D, Świderski A. Laboratory measurements of vehicle exhaust emissions in conditions reproducing real traffic. Measurement 2021; 174: 108998, https://doi.org/10.1016/j.measurement.2021.108998.
10. Coburn T C. Statistical Anaysis and Modeling of Automotive Emissions. Diane Publishing: 2001.
11. Dhahad H A, Hasan A M, Chaichan M T, Kazem H A. Prognostic of diesel engine emissions and performance based on an intelligent technique for nanoparticle additives. Energy 2022; 238: 121855, https://doi.org/10.1016/j.energy.2021.121855.
12. Dimitriadis A, Seljak T, Vihar R et al. Improving PM-NOx trade-off with paraffinic fuels: A study towards diesel engine optimization with HVO. Fuel 2020; 265: 116921, https://doi.org/10.1016/j.fuel.2019.116921.
13. EL-Seesy A I, Kayatas Z, Takayama R et al. Combustion and emission characteristics of RCEM and common rail diesel engine working with diesel fuel and ethanol/hydrous ethanol injected in the intake and exhaust port: Assessment and comparison. Energy Conversion and Management 2020; 205: 112453, https://doi.org/10.1016/j.enconman.2019.112453.
14. Falai A, Misul D A. Data-Driven Model for Real-Time Estimation of NOx in a Heavy-Duty Diesel Engine. Energies 2023; 16(5): 2125, https://doi.org/10.3390/en16052125.
15. Fuc P, Lijewski P, Kurczewski P et al. The Analysis of Fuel Consumption and Exhaust Emissions From Forklifts Fueled by Diesel Fuel and Liquefied Petroleum Gas (LPG) Obtained Under Real Driving Conditions. Volume 6: Energy, Tampa, Florida, USA, American Society

of Mechanical Engineers: 2017: V006T08A060, https://doi.org/10.1115/IMECE2017-70158.
16. Haasz T, Gómez Vilchez J J, Kunze R et al. Perspectives on decarbonizing the transport sector in the EU-28. Energy Strategy Reviews 2018; 20: 124–132, https://doi.org/10.1016/j.esr.2017.12.007.
17. Hosseini M, Chitsaz I. Knock probability determination in a turbocharged gasoline engine through exhaust gas temperature and artificial neural network. Applied Thermal Engineering 2023; 225: 120217, https://doi.org/10.1016/j.applthermaleng.2023.120217.
18. Hosseini S H, Taghizadeh-Alisaraei A, Ghobadian B, Abbaszadeh-Mayvan A. Artificial neural network modeling of performance, emission, and vibration of a CI engine using alumina nano-catalyst added to diesel-biodiesel blends. Renewable Energy 2020; 149: 951–961. https://doi.org/10.1016/j.renene.2019.10.080
19. Jamrozik A, Tutak W, Pyrc M et al. Study on co-combustion of diesel fuel with oxygenated alcohols in a compression ignition dual-fuel engine. Fuel 2018; 221: 329–345, https://doi.org/10.1016/j.fuel.2018.02.098.
20. Karunamurthy K, Janvekar A A, Palaniappan P L et al. Prediction of IC engine performance and emission parameters using machine learning: A review. Journal of Thermal Analysis and Calorimetry 2023; 148(9): 3155–3177, https://doi.org/10.1007/s10973-022-11896-2.
21. Khurana S, Saxena S, Jain S, Dixit A. Predictive modeling of engine emissions using machine learning: A review. Materials Today: Proceedings 2021; 38: 280–284, https://doi.org/10.1016/j.matpr.2020.07.204.
22. Kriaučiūnas D, Žvirblis T, Kilikevičienė K et al. Impact of Simulated Biogas Compositions (CH4 and CO2) on Vibration, Sound Pressure and Performance of a Spark Ignition Engine. Energies 2021; 14(21): 7037, https://doi.org/10.3390/en14217037.
23. Li X-Q, Song L-K, Bai G-C. Deep learning regression-based stratified probabilistic combined cycle fatigue damage evaluation for turbine bladed disks. International Journal of Fatigue 2022; 159: 106812, https://doi.org/10.1016/j.ijfatigue.2022.106812.
24. Li X-Q, Song L-K, Bai G-C, Li D-G. Physics-informed distributed modeling for CCF reliability evaluation of aeroengine rotor systems. International Journal of Fatigue 2023; 167: 107342, https://doi.org/10.1016/j.ijfatigue.2022.107342.
25. Li X-Q, Song L-K, Choy Y-S, Bai G-C. Multivariate ensembles-based hierarchical linkage strategy for system reliability evaluation of aeroengine cooling blades. Aerospace Science and Technology 2023; 138: 108325, https://doi.org/10.1016/j.ast.2023.108325.
26. Lin L, Liu J, Guo H et al. Sample adaptive aero-engine gas-path performance prognostic model modeling method. Knowledge-Based Systems 2021; 224: 107072, https://doi.org/10.1016/j.knosys.2021.107072.
27. Ming W J, Min H, Jun Y, Jun L X. The Study of Process Reliability of Aircraft Engine. Procedia Engineering 2015; 99: 835–839, https://doi.org/10.1016/j.proeng.2014.12.609.
28. Mirhashemi F S, Sadrnia H. NOX emissions of compression ignition engines fueled with various biodiesel blends: A review. Journal of the Energy Institute 2020; 93(1): 129–151, https://doi.org/10.1016/j.joei.2019.04.003.
29. Raghuvaran S, Ashok B, Veluchamy B, Ganesh N. Evaluation of performance and exhaust emission of C.I diesel engine fuel with palm oil biodiesel using an artificial neural network. Materials Today: Proceedings 2021; 37: 1107–1111, https://doi.org/10.1016/j.matpr.2020.06.344.
30. Ramachander J, Gugulothu S, Sastry G R, Surya M S. Statistical and experimental investigation of the influence of fuel injection strategies on CRDI engine assisted CNG dual fuel diesel engine. International Journal of Hydrogen Energy 2021. https://doi.org/10.1016/j.ijhydene.2021.04.010
31. Rimkus A, Matijošius J, Manoj Rayapureddy S. Research of Energy and Ecological Indicators of a Compression Ignition Engine Fuelled with Diesel, Biodiesel (RME-Based) and Isopropanol Fuel Blends. Energies 2020. https://doi.org/10.3390/en13092398
32. Romig C, Spataru A. Emissions and engine performance from blends of soya and canola methyl esters with ARB #2 diesel in a DDC 6V92TA MUI engine. Bioresource Technology 1996; 56(1): 25–34, https://doi.org/10.1016/0960-8524(95)00175-1.
33. Saravanan P, Kumar N M, Ettappan M et al. Effect of exhaust gas re-circulation on performance, emission and combustion characteristics of ethanol-fueled diesel engine. Case Studies in Thermal Engineering 2020; 20: 100643, https://doi.org/10.1016/j.csite.2020.100643.
34. Sevinc H, Hazar H. Investigation of performance and exhaust emissions of a chromium oxide coated diesel engine fueled with dibutyl maleate mixtures by experimental and ANN technique. Fuel 2020; 278: 118338, https://doi.org/10.1016/j.fuel.2020.118338.
35. Sharma P, Sharma A Kr, Balakrishnan D et al. Model-prediction and optimization of the performance of a biodiesel – Producer gas powered dual-fuel engine. Fuel 2023; 348: 128405, https://doi.org/10.1016/j.fuel.2023.128405.
36. Sivaramakrishnan K, Ravikumar P. Optimization of operational parameters on performance and emissions of a diesel engine using

biodiesel. International Journal of Environmental Science and Technology 2014; 11(4): 949–958. https://doi.org/10.1007/s13762-013-0273-5

37. Szabados G, Knaup J, Bereczky Á. ICE Relevant Physical-chemical Properties and Air Pollutant Emission of Renewable Transport Fuels from Different Generations – An Overview. Periodica Polytechnica Transportation Engineering 2022; 50(1): 11–22, https://doi.org/10.3311/PPtr.14925.
38. Szabados G, Nagy P, Zsoldos I, Rohde-Brandenburger J. Comparing the Combustion Process and the Emission Characteristic of a Stationary Heating Device System and an Internal Combustion Engine with Experimental Investigation. Periodica Polytechnica Transportation Engineering 2023; 51(1): 96–104, https://doi.org/10.3311/PPtr.18751.
39. Taghizadeh-Alisaraei A, Ghobadian B, Tavakoli-Hashjin T, Mohtasebi S S. Vibration analysis of a diesel engine using biodiesel and petrodiesel fuel blends. fuel 2012; 102: 414–422. https://doi.org/10.1016/j.fuel.2012.06.109
40. Thurston M G, Sullivan M R, McConky S P. Exhaust-gas temperature model and prognostic feature for diesel engines. Applied Thermal Engineering 2023; 229: 120578, https://doi.org/10.1016/j.applthermaleng.2023.120578.
41. Uludamar E, Tosun E, Aydın K. Experimental and regression analysis of noise and vibration of a compression ignition engine fuelled with various biodiesels. Fuel 2016; 177: 326–333. https://doi.org/10.1016/j.fuel.2016.03.028
42. Warguła Ł, Lijewski P, Kukla M. Correction to: Influence of non-commercial fuel supply systems on small engine SI exhaust emissions in relation to European approval regulations. Environmental Science and Pollution Research 2022; 29(37): 55944–55944, https://doi.org/10.1007/s11356-022-20372-1.
43. Zhang Y, Wang Q, Chen X et al. The Prediction of Spark-Ignition Engine Performance and Emissions Based on the SVR Algorithm. Processes 2022; 10(2): 312, https://doi.org/10.3390/pr10020312.
44. Zöldy M. ETHANOL–BIODIESEL–DIESEL BLENDS AS A DIESEL EXTENDER OPTION ON COMPRESSION IGNITION ENGINES. TRANSPORT 2011; 26(3): 303–309, https://doi.org/10.3846/16484142.2011.623824.
45. Zöldy M. Fuel properties of butanol–hydrogenated vegetable oil blends as a diesel extender option for internal combustion engines. Periodica Polytechnica Chemical Engineering 2020; 64(2): 205–212. https://doi.org/10.3311/PPch.14153
46. Žvirblis T, Vainorius D, Matijošius J et al. Engine Vibration Data Increases Prognosis Accuracy on Emission Loads: A Novel Statistical Regressions Algorithm Approach for Vibration Analysis in Time Domain. Symmetry 2021; 13(7): 1234, https://doi.org/10.3390/sym13071234.
47. Ecology in transport: problems and solutions. Cham, Springer: 2021.