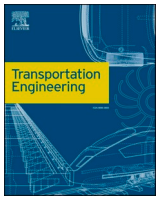
Transportation Engineering 20 (2025) 100316

Contents lists available at ScienceDirect

Transportation Engineering

journal homepage: www.sciencedirect.com/journal/transportation-engineering



Machine learning assisted modeling of ignition delay in a light-duty gasoline compression ignition engine

Rossi Alessandro, Silvagni Giacomo *, Ravaglioli Vittorio, Corti Enrico

DIN – Dipartimento di Ingegneria Industriale, Alma Mater Studiorum – Università di Bologna, Bologna, 40121 Italy

ARTICLE INFO

Keywords:
Ignition delay modeling
Artificial neural network
Sustainable fuels and powertrains
Gasoline compression ignition
Control-oriented strategy

ABSTRACT

This work proposes an innovative artificial neural network-based (ANN) approach to predict the ignition delay (ID) in a Gasoline Compression Ignition (GCI) engine using the information coming from standard sensors mounted on the engine. Moving toward the carbon neutrality of transport by using renewable and synthetic fuels, GCI combustion is considered a promising technology to achieve high engine efficiency and ultra-low pollutant emissions. As with other auto-ignition combustion concepts, a deep understanding of compression ignition dynamics is crucial for maintaining a stable and controllable combustion process in different operating and environmental conditions and enhancing engine performance and durability over time. Machine learning (ML) offers a promising modelling tool to lower the cost for testing control strategies compared to traditional physical or empirical approaches. An experimental campaign was conducted with 293 steady-state engine operating conditions to establish relationships between ignition delay and engine parameters, such as engine speed, load, intake and injection pressures, exhaust gas recirculation (EGR), and injection parameters, in a light-duty GCI engine. Then, an ANN-based model was trained and validated using a combination of holdout and k-fold cross-validation methods, along with a Bayesian regularization algorithm. The ignition delay estimation through the ANN-based model has shown a NRMSE percentage of 2.16 % and R^2 of 0.99 on test dataset, demonstrating to be accurate enough for engine control and diagnosis purposes. This work aims at inspiring innovative ANN-based control strategies for promoting the use of advanced combustion methodologies, such as GCI combustion, as technical solution for production engines.

1. Introduction

In recent years, there has been an increased push to reduce health-damaging pollutant emissions from road vehicles [1,2]. As shown in Tables 1 and 2, in the European Union (EU) the limits (from Euro 1 to Euro 6-d) of the vehicle emissions powered by spark ignited (SI) engines and compression ignited (CI) engines have decreased considerably [3,4].

While, at the beginning, the restrictions were mainly focused on pollutants such as carbon monoxide (CO), unburned hydrocarbons (UHC), nitrogen oxides (NOx), particulate matter (PM) and particle number (PN), nowadays climate change has become an important topic for global society [5,6]. In order to meet the requirements for the management of greenhouse gas (GHG) emissions, established by Kyoto Protocol [5] and the Paris Agreements [6], the EU has set a stringent Carbon Dioxide (CO2) standard [7] for the new passenger cars from 2021 (95 g/km) [8]. Later, with the objective of a progressive process towards a zero net greenhouse gas emissions economy in EU by 2050 [9], the latest CO2 limits [10] have been reduced by an additional 15 % from 2025 to 37.5 % from 2030. Along with the reduction of limits on CO2 and chemical pollutant species, a change in the homologation cycles has also been introduced. Indeed, in order to monitor the real vehicle emissions and reduce the gap between test bench testing and real driving [11], together with the standard test bench homologation cycles (New European Driving Cycle (NEDC) [12] and World-wide harmonized Light duty Test Procedure (WLTP) [12,13]) the Real Driving Emissions (RDE) test has also been introduced [4,11,14]. As previously mentioned, the RDE cycle must be performed on road and, although the legislation [15] defines a set of guidelines for assessing the validity of the test carried out, the number of possible cycles is unlimited [4]. Indeed, the RDE test objective is to push automotive manufacturers to comply with the limits on pollutant emissions and CO2 production in a wide range of boundary conditions, such as temperature, altitude, driving style and traffic [16]. Ultimately, the stricter limits on pollutant emissions, the introduction of real driving emission test and the need to reduce fuel

* Corresponding authors.
E-mail address: giacomo.silvagni2@unibo.it (S. Giacomo).

https://doi.org/10.1016/j.treng.2025.100316
Received 15 January 2024; Received in revised form 25 January 2025; Accepted 2 March 2025
Available online 4 March 2025

Symbols and abbreviations

AFR	Air-Fuel Ratio
AFR_{st}	Stoichiometric Air-Fuel Ratio
AI	Artificial Intelligence
ANN	Artificial Neural Network
BEV	Battery Electric Vehicle
BOA	Bayesian Optimization Algorithm
BP	Backpropagation algorithm
CA	Crank angle degrees
CAN	Communication Area Network
CA50	Crank angle at which the 50 % of fuel mass inside the cylinder is burnt
CFD	Computational Fluid Dynamics
CIDI	Compression Ignition Direct Injection
CI	Compression Ignition Engines
CO	Carbon Monoxide
CO2	Carbon Dioxide
cp	Specific Heat at Constant Pressure
cv	Specific Heat at Constant Volume
dp	Derivative of In Cylinder Pressure
$d\theta$	Derivative of Crankshaft Angle
DISI	Direct Injection Spark Ignition
dV	Derivative of Combustion Chamber Volume
ECU	Electronic Control Unit
EGR	External Residual Gases
EU	European Union
f	is the nonlinear activation function of the neuron in the hidden layer
FNN	Feedforward Neural Network
g	is the activation function of the output layer
GCI	Gasoline Compression Ignition
GHG	Greenhouse Gas
H	Hessian Matrix
h_j	is the output of the *jth* hidden neuron
HCCI	Homogeneous Charge Compression Ignition
HEV	Hybrid Electric Vehicle
I	Identity Matrix
Iw	Input Weights
Iw_{ji}	is the weight of the connection linking the *ith* neuron in the input layer to the *jth* neuron in the hidden layer
ICE	Internal Combustion Engine
IMEP	Indicating Mean Effective Pressure
IVC	Intake Valve Closing angle
J	Jacobian Matrix
k_{EGR}	Magnitude index of external residual gases in the combustion chamber
k_{iRG}	Magnitude index of internal residual gases in the combustion chamber
km	kilometers
LTC	Low Temperature Combustion
Lb	matrix of the input weights with the addition of biases for *m* neuron
Lb_j	is the bias of the *jth* hidden unit
Lw	matrix of the layer weights with the addition of biases for *k* neuron in the output layer
Lw_{kj}	is the weight of the connection linking the *jth* neuron in the hidden layer to the *kth* neuron in the output layer
M_{EGR}	Mass of external residual gases in the combustion chamber
M_{H_2O}	Mass of water vapor in the combustion chamber
M_{RG}	Mass of internal residuals in the combustion chamber
M_{O_2}	Mass of oxygen of the air in the combustion chamber
M_{eq}	Equivalent Mass
m'	is the number of neuron in the hidden layer
m	is the epoch of the training process
ML	Machine Learning
mg	milligrams
MSE	Mean Squared Error
n	is the number of inputs of the net
N_{input}	Number of input
N_{sample}	Number of samples
NEDC	New European Driving Cycle
NOx	Nitrogen Oxides
NMHC	Non-Methane Hydrocarbons
NRMSE	Percentage normalized root mean square error
Ob	Output Bias
Ob_k	is the bias of the *kth* output unit
P_{IVC}	Combustion Chamber pressure at the IVC
P_{rail}	Rail Pressure
P_{SOI}	Combustion Chamber pressure at the SOI of the pilot injection
p	In-Cylinder Pressure
PE	Percentage Error
PHEV	Plug-in Hybrid Electric Vehicle
PM	Particulate Matter
PN	Particulate Number
R^2	R-squared
R_{air}	Air Constant
RDE	Real Driving Emissions
ReLu	Rectified Linear Unit activation function
RI	Relative Input Importance
RMSE	Root mean square error
$RMSE_i$	Root mean square error of the i th variable
RNN	Recurrent Neural Network
ROHR	Rate of Heat Release
RON	Research Octane Number
SI	Spark Ignition Engines
SOC	Start of Combustion
SOI	Start of Injection
SOI_{pilot}	Start of Injection of pilot injection
T_{cyl}	Bulk gas temperature
T_{SOI}	Temperature at SOI
trainBR	Bayesian regularization backpropagation algorithm
trainGD	Gradient descent algorithm
trainGDM	Gradient descent with momentum algorithm
trainGDX	Gradient descent with momentum and adaptive learning rate
trainLM	Levenberg-Marquardt backpropagation algorithm
trainSCG	Scaled conjugate gradient backpropagation algorithm
tanh	Hyperbolic Tangent activation function
UHC	Unburned Hydrocarbons
V	Combustion Chamber Volume
V_{IVC}	Combustion Chamber Volume at the intake valve closing
V_{SOI}	Combustion Chamber Volume at SOI of the pilot injection
VGT	Variable Geometry Turbine
w	Vector space of weights and bias
WLTP	World-wide harmonized Light duty Test Procedure
X	Vector of the *n* input feature and the additional input for the bias equal to 1
X_{max}	Maximum value of the variable
X_{min}	Minimum value of the variable
X_{norm}	Normalized value of the variable
x_i	is the *ith* input to the hidden unit
$AptCommand0177_{;k}$	refers to the outcome of the *kth* output unit
y_p	Value of the output unit of the FNN
$AptCommand0177_{;p}$	Average value of the output unit of the FNN
α	Regularization term
$\Delta\theta_{ID}$	Ignition Delay in crankshaft angle degree

φ	Equivalence Ratio	σ	Standard deviation of the variable
γ	Specific Heat Ratio	θ_{SOC}	Crankshaft angle at Start of Combustion
μ	Damping parameter	θ_{SOI}	Crankshaft angle at Start of Injection
η	Learning rate,	θ_{ROHR_5}	Crankshaft angle in which ROHR is equal to 5 J/deg
π	Mean of the variable	$\theta_{SOI_{pilot1}}$	Crankshaft angle of the first Pilot Start of Injection

consumption (CO2 reduction) have pushed the automotive industry towards greener technologies.

Although vehicle electrification seems to be the most promising option to drastically reduce the use of fossil fuels [17,18], the complete conversion of the transport sector from traditional powertrains (SI and DI) to fully electric vehicles would require a re-planning of the energy supply system of many countries [19]. In addition to this, Battery Electric Vehicles (BEV) are facing specific challenges such as: short operating range, long recharging time and high bulk and weight [20]. The research community and the industry are currently focused on solving these problems for the mobility of the future, but the transition will be progressive and hybrid systems [21,22] are considered crucial. For these reasons, the combustion engine will be present on the market for at least another 13 years, most likely in Hybrid Electric Vehicle (HEV).

In this scenario, the effort of the researchers on Internal Combustion Engines (ICE) has been redirected towards the study and development of high efficiency combustion methodologies, commonly named Low Temperature Combustion (LTC), that can be used in HEV or Plug-in Hybrid Electric Vehicle (PHEV) to further reduce CO2 production.

LTC techniques are identified by the spontaneous ignition of highly diluted air-fuel blends. The high compression ratios, the short combustion duration, and the low pumping losses, make these types of combustion much more efficient compared to traditional combustion processes [23]. Moreover, thanks to the low local temperature and the high dilution of the air-fuel mixture inside the combustion chamber, LTCs have significantly lower NOx and PM emissions than Compression Ignition Direct Injection (CIDI) and Direct Injection Spark Ignition (DISI) engines [23–26]. Notwithstanding the significant benefits of LTC combustion strategies, to close the gap between trial applications and industrial needs, there are still limitations and challenges that need to be addressed. Limited operating range, higher hydrocarbon and carbon monoxide emissions, sensitivity to fuel properties, higher noise and vibration and high cycle-to-cycle variability [27] represent the main limitations for a wide diffusion of LTC. Furthermore, since the ignition process of these advanced combustion processes is driven by chemical reactions, the preparation of the fuel-air mixture is a crucial factor that can potentially compromise the stability of the combustion [28].

Gasoline Compression Ignition (GCI) combustion has emerged as one of the most promising LTC concepts for achieving high-efficiency combustion in ICEs. GCI combustion involves the high-pressure direct injection of gasoline during the compression stroke, which leads to the creation of a global lean mixture. Once the ignition delay has elapsed, the air-fuel mixture spontaneously releases energy, allowing for precise control of the combustion process [24–26,29,30]. Although the potential of GCI engines has been widely demonstrated, their application is limited by knock and misfire, which can compromise engine reliability,

Table 2
Limits for pollutant emissions from Euro 1 to Euro 6-d for CI engines.

		Date	CO [mg/km]	HC + NOx [mg/km]	NOx [mg/km]	PM [mg/km]	PN [#/km]
Compression Ignited Engines (CI)	*Euro 1*	*1992*	2720	970	–	140	–
	Euro 2	*1996*	1000	700	–	80	–
	Euro 3	*2000*	640	560	500	50	–
	Euro 4	*2005*	500	300	250	25	–
	Euro 5-b	*2011*	500	230	180	4.5	6×10^{11}
	Euro 6-d	***2020***	**500**	**170**	**80**	**4.5**	$\mathbf{6 \times 10^{11}}$

and by a narrow operating range [31]. However, several studies have shown the benefit of multiple fuel injections to enhance the stability of the combustion process and extend the GCI engine operating range [32,33]. Wei et al. [31] demonstrated that GCI combustion controllability heavily depends on the heat released by a small fuel quantity introduced through pilot injections. As a matter of fact, the increase in-cylinder pressure and temperature caused by pilot injections impact the auto-ignition processes, leading to decrease the ignition delay of the subsequent injection. Such modifications in the thermodynamic of the air-fuel charge promotes a smoother combustion process, and the following injection burns in stratified conditions [34,35]. Despite the remarkable benefits related to the use of multiple injection strategies in GCI [25,26], wrong placement of pre-injections adversely affects combustion stability, leading to the occurrence of misfire or knock [36,37]. Moreover, the use of gasoline-like fuels (with a reactivity lower than diesel) can result in a longer ignition delay (ID) of the first injection, which might compromise the stability of the whole combustion, especially in cold operating conditions [38] and at low load/speed. For this reason, innovative injection controllers, based on accurate ID estimation and able to dynamically adjust the injection pattern, become essential to extend the GCI operating range and maximize its benefits both in terms of efficiency and pollutants, especially at low loads and speed.

Previous studies have shown that a proper selection of pilot injections pattern, i.e., number of injections, angular position, and injected mass, can ensure the GCI stability across a wide operating range [37]. However, due to the need for a large experimental activity to test GCI in various operating conditions, the standard calibration process can be very expensive, time-consuming, and hardly generalizable to

Table 1
Limits for pollutant emissions from Euro 1 to Euro 6-d for SI engines.

		Date	CO [mg/km]	HC + NOx [mg/km]	UHC [mg/km]	NOx [mg/km]	NMHC [mg/km]	PM [mg/km]	PN [#/km]
Spark Ignited Engines (SI)	*Euro 1*	*1992*	2720	970	–	–	–	–	–
	Euro 2	*1996*	2200	500	–	–	–	–	–
	Euro 3	*2000*	2300	–	200	150	–	–	–
	Euro 4	*2005*	1000	–	100	80	–	–	–
	Euro 5-b	*2011*	1000	–	100	60	68	4.5	–
	Euro 6-d	***2020***	**1000**	**–**	**100**	**60**	**68**	**4.5**	$\mathbf{6 \times 10^{11}}$

unknown operating conditions.

In recent years, model-based approaches have emerged as the most promising solution to reduce test time and costs required for calibration of the engines. Despite computational fluid dynamics (CFD) models are capable of providing accurate predictions of the ID, their use in real-time control is limited by their high computational cost [39]. Empirical ID models based on the Arrhenius equation have been proposed in the literature as a computationally efficient alternative for developing control models for combustion engines [39,40]. However, empirical models may not always capture the complexities of GCI combustion dynamics mainly because such models assume a homogenous fuel-air mixture, while the fuel-air mixture in GCI engines can be highly heterogeneous, leading to significant variations in ID. To overcome this limitation, Silvagni et al. [41] proposed a control-oriented ignition delay model based on physical parameters, which summarize the chemical-physical and thermodynamic properties of the air-fuel mixture. Despite the model has demonstrated a good accuracy on the ID prediction in different engine conditions, it does not take into account the effects of fuel injection pressure and engine speed, which are known having a significant impact on ignition delay in GCI engines.

Meanwhile, the application of machine learning (ML) methodologies in engine fields is showing a high potential to overcome the limitations of CFD models) [42–45]. Nowadays, artificial neural networks (ANN) have emerged as a powerful tool for predicting, analyzing, and optimizing a wide range of characteristics associated with internal combustion engines, including performance, combustion phasing, thermodynamics and chemical interactions in the air-fuel mixture, and emissions levels across a variety of combustion technologies (gasoline, diesel, and HCCI) [42–45]. Many studies in the literature have shown that ML approaches can achieve high accuracy in modeling complex phenomena with more input variables than traditional physical models in control applications. Despite the advantages described above, the quality and representativeness of the training dataset, and the input features used in ML algorithms can greatly affect the accuracy, precision, and generalizability of the model [46]. For these reasons, the model selection process is extremely important in machine learning applications, and it must be carefully considered [47].

This study presents an ID model using a machine learning approach aimed at improving the performance and controllability of a light-duty GCI engine. The ID is modeled using a feedforward neural network (FNN) based on multiple inputs (selected based on the physics of the system) that can be directly measured with standard sensors already present in the stock engine. To obtain the database for the ANN training and testing, an experimental activity was carried out running a light-duty 1.3 L turbocharged CI engine converted to run GCI combustion. Various training algorithms and FNN configurations were evaluated using well-known holdout [48] and k-fold cross-validation [49] methods to determine the optimal model for the ID estimation [50]. Then, using the information coming from the experimental campaign, the ANN ID model was validated showing good accordance with the measured ID.

2. Experimental setup

As previously stated, with the aim of extending the operating range of the GCI engines through the optimization of the fuel injection pattern, this study proposes a control-oriented ID model using an FNN. This section outlines the experimental setup used to test a light-duty GCI engine on a test bench. Such engine platform was used to produce the dataset needed for training and testing the neural network model. The technical specifications of the engine are presented in Table 3.

To realize GCI combustion, the stock fuel injection system of the engine (a standard common-rail high pressure injection system), was fueled with commercial RON 95 gasoline. Thanks to the capability of the injection system of performing multiple injections by using specifically designed patterns, an accurate control of the combustion process was achieved. Several works in literature pointed out the benefits of pilot injections for improving GCI stability and controllability. As a result, during the experimental activity, the selected injection pattern was composed of two pilot injections, followed by the main injection.

Table 3
Engine technical characteristics.

Engine Parameter	Values
Engine type	L4, four-stroke
Displaced volume	1248 cc
Maximum Torque	200 Nm @ 1500 rpm
Maximum Power	70 kW @ 3800 rpm
Injection System	Common Rail, Multi-Jet
Bore	69.6 mm
Stroke	82 mm
Compression ratio	16.8:1
Firing Order	1–3–4–2

In order to maximize the benefits of GCI in terms of NOx and PM emissions, a high-pressure exhaust gas recirculation (EGR) system was installed. Additionally, to properly manage the intake pressure, the engine was equipped with a variable geometry turbine actuator (VGT).

To promote gasoline autoignition, particularly during cranking and cold start conditions, when pressure and temperature of the inlet air are critically low, external devices were added to the stock engine: a roots volumetric compressor (S/C Eaton Compressor M24), and a diathermic oil thermoregulation unit (TEMPCO T-REG HCE 609/15-O). As shown in Fig. 1, the Eaton compressor was installed upstream the standard dynamic compressor. An electric motor (5.5 kW and maximum rotational speed of 3000 rpm) drives the volumetric compressor, which is turned off once the cranking phase is completed. Thereafter, the boost pressure is directly controlled through the standard turbocharger using the VGT.

As is commonly known, effective management of GCI combustion requires control strategies that differ from those traditionally employed in standard engine control units (ECU), i.e., model-based control strategies. Accurate control of engine actuators such as VGT, EGR valve, and injection system as well as auxiliary systems using unconventional control strategies is essential to maximize the potential of GCI combustion. To enhance the flexibility of the engine control system and overcome the limitations of the standard control unit in terms of customization, the stock ECU was replaced with a fully programmable ECU (Spark by Alma Automotive), based on National Instruments hardware and programmable via LabView software. To perform an in-depth analysis of GCI combustion, four in-cylinder pressure sensors (AVL GH14P), one for each cylinder, were mounted. Slow-speed sensors, such as intake and exhaust pressure and temperature sensors, were acquired by the test bed acquisition system and the ECU, while the cylinder pressure signals were acquired, at a rate of 200 kHz, by an external combustion analysis system (OBI by Alma Automotive).

The combustion metrics calculated in real-time by the combustion analyzer were sent to the Open ECU at high-speed using a controller area network (CAN) bus. This allowed for a data exchange at a frequency greater than the combustion frequencies tested during the experimental campaign. Consequently, the combustion indexes were used as inputs for cycle-to-cycle closed-loop combustion controllers implemented in the Open ECU. Throughout the whole experimental activity, thanks to the closed-loop control strategy the indicated mean effective pressure (IMEP) and center of combustion (CA50) delivered by the engine were kept at target values. Fig. 1 shows the GCI engine layout and its integration with the testing facilities.

In order to select the best ANN model and evaluate its predictive capabilities, a total of 293 steady-state engine operating conditions were investigated. Table 4 provides the range of values for the main engine and external parameters in the dataset. It is worth mentioning that, to obtain a reliable database for ID modelling, the GCI combustion was tested only in stable and safe conditions for the engine, i.e. where the coefficient of variation of IMEP was kept below 5 %, maximum in-cylinder pressure below 140 bar, and maximum peak pressure rise

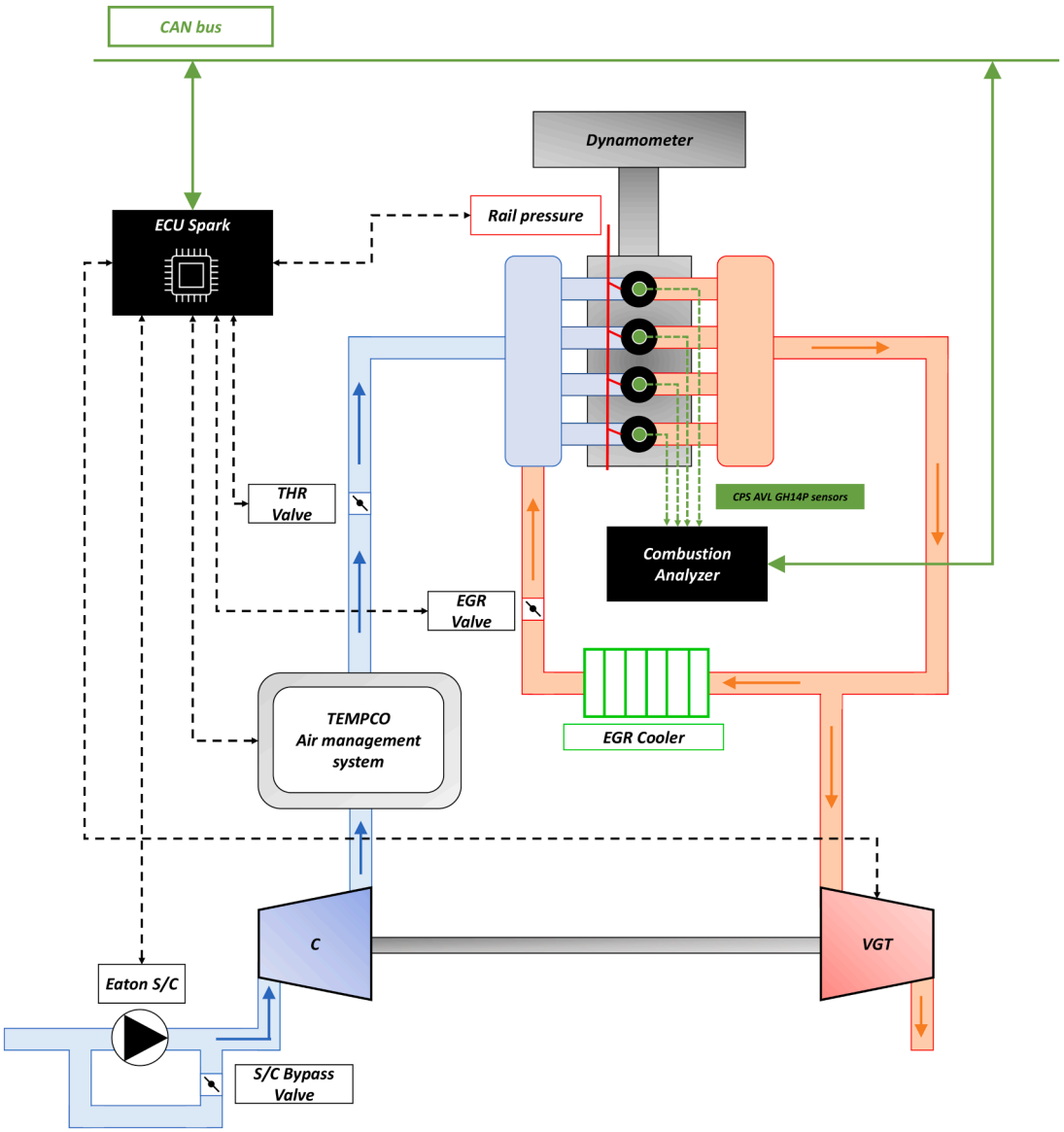


Fig. 1. GCI engine experimental layout.

rate below 12 bar/deg. Since GCI combustion is characterized by high combustion instability, low loads and speeds were excluded from ANN ID model training and validation.

Table 4
Minimum and maximum values of the main engine and external variables on dataset.

Engine parameter	Values
Engine speed [rpm]	2000 – 3500
IMEP [bar]	8 – 14
CA50 [ca. deg aTDC]	6 – 18
Injection pressure [bar]	350 – 800
SOI pilot [deg bTDC]	32 – 62
SOI pre [deg bTDC]	18 – 39
SOI main [deg bTDC]	−1 – 14
Boost pressure [bar]	1.5 – 2.2
Intake temperature [°C]	37 – 78
Lambda [-]	1 – 2.6
Exhaust pressure [bar]	1.8 – 3
ET main [μs]	350 - 1100
EGR rate [%]	0 – 22
Relative umidity [%]	19 – 52

It is worth noting that the FNN model presented in this work was designed and validated considering only on one cylinder (cylinder 3) of the engine. Despite this limitation, after retraining the FNN, the methodology can be extended to other cylinders of the same engine without any modification.

3. Feedforward neural network (FNN) for supervised learning

In the field of machine learning (ML), two fundamental approaches are widely used: supervised learning and unsupervised learning. Supervised learning involves training an algorithm on a labeled dataset, in which each data point is associated with a known target or outcome variable. This approach is commonly used in regression and classification problems. In contrast, unsupervised learning involves training an algorithm on an unlabeled dataset, in which there are no known outcomes or target variables. The algorithm will identify patterns and relationships in the database, relying exclusively on the provided input features. Unsupervised learning is often used for tasks such as clustering, anomaly detection, and dimensionality reduction [51,52]. In this study, to address a regression problem, a supervised learning approach was chosen. The goal was to train a model, based on FNN architecture, to

predict a continuous target variable based on a set of input features.

Feedforward neural networks (FNN) have gained significant attention across various fields owing to their flexibility and remarkable representation capabilities. In supervised learning these networks can capture complex and nonlinear relationships between input and output data making them a valuable tool in several applications, such as classification and regression [53]. Unlike recurrent neural network (RNN), feedforward neural networks do not have loops in their architecture and do not store past information. This one-directional data flow makes FNN well-suited for tasks where the current input is the only relevant information.

A fully connected FNN typically consists of an input layer (one or multiple hidden layers) and an output layer. Each hidden layer can have multiple neurons, and every neuron in each layer is connected to nodes in the next layer. The fully connected architecture ensures that each neuron in each layer is fully connected to all the neurons in the previous and subsequent layers. Fig. 2 illustrates the structure of a fully connected FNN for supervise learning.

The schematic shows an input layer with input weights (*Iw*), one hidden layer with layer weights (*Lw*) and bias (*Lb*), and an output layer with *k* output variables. The input layer receives *n* input features and passes them to the first hidden layer, which consists of *m* neurons fully connected to the input layer. The output of each hidden unit is passed through an activation function (*f*) and then sent to the output layer. The output layer contains *k* neuron (one for each output variables) with bias (*Ob*) and their own activation function (*g*), which compute the final outputs of the network ($\hat{y}$).

The output of an FNN can be calculated mathematically using the following formulas:

$$h_j = f\left(\sum_{i=1}^{n} x_i Iw_{ji} + Lb_j\right) \tag{1}$$

$$AptCommand0177_{;k} = g\left(\sum_{j=1}^{m'} h_j Lw_{kj} + Ob_k\right) \tag{2}$$

Where:

- n is the number of inputs unit;
- h_j is the output of the *jth* hidden neuron;
- x_i is the *ith* input to the hidden unit;
- Iw_{ji} is the weight of the connection linking the *ith* neuron in the input layer to the *jth* neuron in the hidden layer;
- Lb_j is the bias of the *jth* hidden unit;
- f is the nonlinear activation function of the neuron in the hidden layer (typically sigmoid or hyperbolic tangent (tanh) or rectified linear unit (ReLu)) [54];
- f refers to the outcome of the *kth* output unit;
- m' is the number of neuron in the hidden layer;
- Lw_{kj} is the weight of the connection linking the *jth* neuron in the hidden layer to the *kth* neuron in the output layer;
- Ob_k is the bias of the *kth* output unit;
- g is the activation function of the output layer (typically linear for regression problem), which can be different from that of the hidden layer;

Finally, by substituting Eq. (1) into Eq. (2), it is possible to obtain the general expression that relates the *n* input variables to the *kth* output variable (Eq. (3)):

$$AptCommand0177_{;k} = g\left(\sum_{j=1}^{m} Lw_{kj} f\left(\sum_{i=1}^{n} x_i Iw_{ji} + Lb_j\right) + Ob_k\right) \tag{3}$$

In order to represent the output of a FNN in a matrix form, the biases of the hidden and output layers can be treated as additional units. Specifically, the bias term Lb_j for each neuron *j* in the hidden layer can be considered as an additional input unit with a fixed input value of 1, while the bias term Ob_k for each output neuron *k* can be considered as an additional hidden unit with a fixed output value of 1. By incorporating these additional units, the output of an FNN with input feature vector $\boldsymbol{X} \in \mathbb{R}^{n+1,1}$ can be expressed as (Eq. (4)):

$$AptCommand0177_{;k}(X) = g(LWf(IWX)) \tag{4}$$

Where:

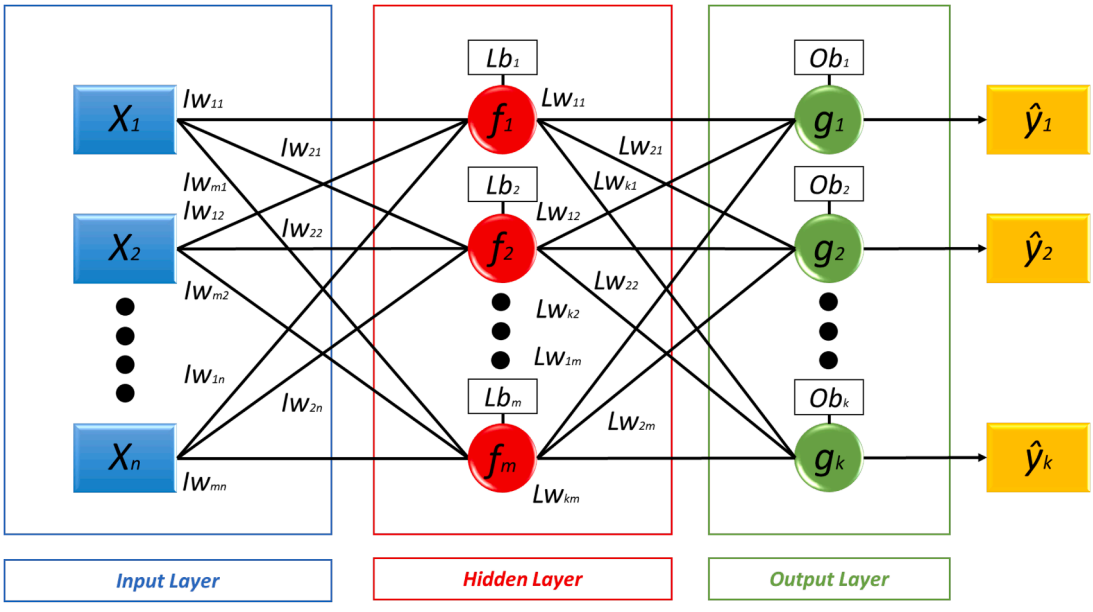


Fig. 2. Structure of feedforward neural network with one hidden layer and one output.

- $X = \begin{bmatrix} 1 \\ x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix}$ is the vector of the n input feature and the additional input for the bias equal to 1;
- $IW = \begin{bmatrix} 1 & \dots & 0 \\ Lb_1 & \dots & Iw_{1n} \\ \vdots & \vdots & \vdots \\ Lb_m & \cdots & Iw_{mn} \end{bmatrix} \in \mathbb{R}^{m+1,\ n+1}$ is the matrix of the input weights with the addition of biases for m neuron in the hidden layer;
- $LW = \begin{bmatrix} Ob_1 & \dots & Lw_{1m} \\ Ob_2 & \dots & Lw_{2m} \\ \vdots & \vdots & \vdots \\ Ob_k & \cdots & Lw_{km} \end{bmatrix} \in \mathbb{R}^{k,\ m+1}$ is the matrix of the layer weights with the addition of biases for k neuron in the output layer;

Despite the universal approximation theorem proposed by Kolmogorov [55] provides a strong theoretical explanation for neural computation, the theorem does not provide guidance on how to compute the appropriate weights, and bias. As a results, the main issue in neural network model implementation for supervised learning is the optimal selection of weights and biases based on training pattern and target output.

3.1. Backpropagation (BP) algorithm

The backpropagation algorithm (BP) is an extremely common approach for determining the optimal weights and biases in a neural network model. BP Algorithm typically begins by initializing the weights and biases of an untrained neural network with random values. Then, the algorithm iteratively minimizes a cost function which depends on the initial weights and biases of the network.

Two distinct methods can be employed to implement the BP algorithm: batch mode and incremental (or online) mode. In batch mode, the entire dataset is presented to the training algorithm during each epoch, and weights and biases are updated based on the average of the descent gradients computed across the entire dataset. Conversely, in incremental mode, the training data is fed to the neural network one by one, or in small groups called mini-batches, during each iteration. Then, weights and biases are updated after processing each example. Although both methodologies have their own advantages and disadvantages, considering the relatively small size of the dataset and the enhanced stability it provides in achieving a solution, the batch mode was chosen as the preferential approach for this study [56]. Therefore, all subsequent descriptions and analyses will be conducted within the context of the batch mode.

The most common cost function used for BP learning is the mean square error (MSE), expressed by Eq. (5), and calculated as the sum of the squared differences between the target value of the output and the output of the neural network.

$$MSE = \frac{1}{N_{sample}} \sum_{p=1}^{N_{sample}} \left(y_p - AptCommand0177_{;p} \right)^2 = E \tag{5}$$

During the training process, the weights and biases are updated by BP algorithm according to specific update rules, summarized by Eq. (6) and Eq. (7), where η is the learning rate, w is the vector space of weights and bias, and m is the epoch of the training process.

$$\Delta w = -\eta \frac{\delta E(m)}{\delta w} \tag{6}$$

$$w(m+1) = w(m) + \Delta w(m) \tag{7}$$

To avoid the presence of plateaus or multiple minima in the error surfaces, an additional term called momentum is often incorporated into the weight update rule. By including a fraction of the previous weight, this term fastens the learning process and dampen oscillations in the solution. Therefore, the weight update rule is modified according to Eq. (8), where m and $m-1$ are the m iteration and the $m-1$ iteration of the training process respectively, and α represents the momentum.

$$\Delta w(m) = -\eta \frac{\delta E(m)}{\delta w} + \alpha \Delta w(m-1) \tag{8}$$

In addition to the standard BP algorithms, i.e., gradient descent (trainGD) and gradient descent with momentum (trainGDM), to reduce the influence of the learning rate on solution stability and convergence, an alternative algorithm called gradient descent with momentum and adaptive learning rate (trainGDX) was developed [57]. To further enhance the convergence of neural network models, a class of algorithms known as "second-order methods" has been introduced. These methods consider not only the first derivative of the error function but also the second derivative with respect to the model parameters. By leveraging the Hessian matrix, which provides information about the curvature of the error curve, these methods reduce the likelihood of being trapped in local minima and facilitate the discovery of better overall solutions. However, it is important to note that second-order methods come with certain drawbacks. They typically involve higher computational costs due to the calculation and inversion of the Hessian matrix, which its existence may not be always guaranteed. To address the challenges associated with second-order methods, various algorithms were developed. Among those, the Levenberg-Marquardt BP algorithm (trainLM), the Scaled conjugate gradient BP algorithm (trainSCG) and the Bayesian regularization BP algorithm (trainBR) are widely utilized. Those algorithms offer distinct approaches to enhance convergence and weight adaptation during the neural network training process.

The Levenberg-Marquardt algorithm combines the benefits of both trainGD and Gauss-Newton methods by including a damping parameter (μ) that adjusts the step size during weight updates [58]. In the trainLR algorithm, the approximated Hessian matrix (H), as represented in Eq. (9), is derived from the Jacobian (J) matrix, and the weight update rule as given by Eq. (10), incorporates I and J.

$$H = J^T J \tag{9}$$

$$\boldsymbol{w}(m+1) = \boldsymbol{w}(m) - \left[J^T J + \mu I \right]^{-1} J^T E \tag{10}$$

The Scaled conjugate gradient algorithm is based on conjugate gradient optimization, but it considers scaling factors to adapt the step size dynamically. This allows for efficient convergence even in the presence of varying sensitivities and irregular error surfaces. In the Bayesian regularization algorithm, the weight update rule follows the Levenberg-Marquardt algorithm, but it integrates a regularization term (E_w) in the error function. Such term, defined by Eq. (11), represents the sum of squared of the network weights and mitigates overfitting by penalizing large weights in the neural network. The term α, defined as Bayes' rule, limits the impact of the regularization term on E_T.

$$E_T = E + \alpha E_w \tag{11}$$

The above presented algorithms offer effective strategies to overcome the limitations associated with second-order methods, while still maintaining accuracy and stability in the training process. Moreover, those algorithms not only impact the training phase but also influence the predictive capabilities of the resulting model [59]. As a result, the selection of the training algorithm plays a crucial role in defining the best neural network model for each purpose.

3.2. Input normalization methods

When working with neural networks, it is crucial to normalize the input variables. The normalization process ensures that variables with different magnitudes are properly scaled to a similar range. This is important because variables with larger values might dominate the training process, causing numerical instabilities and biased learning.

There are a variety of methods of input normalization, the most common are: min-max method, and z-score method. In the min-max method, as described in Eq. (12), where X represents the original value of the variable, X_{norm} is the normalized value, X_{min} is the minimum value of the variable, X_{max} is the maximum value of the variable, and a and b are the desired limits for the $X_{norm.}$, each variable is typically rescaled to a range between −1 and 1 or 0 and 1.

$$X_{norm} = a + \frac{(b-a)*(X - X_{min})}{X_{max} - X_{min}} \tag{12}$$

In the z-score method, defined by Eq. (13), where X represents the original value of the variable, X_{norm} is the normalized value, π is the mean of the variable, and σ is the standard deviation of the variable, all input variables are scaled in a way of having null mean value and a standard deviation of 1.

$$X_{norm} = \frac{X - \pi}{\sigma} \tag{13}$$

Overall, the normalization of input variables is an important step in preparing data for neural network training because, as demonstrated in literature [60], it can yield different results and affect the performance and stability of the developed neural network model.

3.3. Model generalization and parameters optimization

3.3.1. Neural network model generalization

To maximize the accuracy and the reliability of ANN models, it is crucial to address the issue of overfitting. One common approach involves incorporating a validation dataset during the training phase. By evaluating the performance of the model on an independent dataset, the risk of overfitting can be mitigated. The training process ends when the cost function reaches its minimum on the validation dataset. This approach ensures that the model captures the essential patterns necessary for accurate predictions on unseen data.

Another technique that proves beneficial in reducing overfitting is regularization, which increases the magnitude of the cost function with a regularization term, as represented by Eq. (11). By using α, the model promotes simpler solutions and decreases the influence of noisy or irrelevant features. It is important to note that regularization introduces an additional parameter that must be properly chosen to obtain maximum accuracy from neural network model.

As previously mentioned, a widely adopted technique to address overfitting is dataset division into training, validation, and test subsets. This approach involves resampling the dataset to create separate subsets for model training, performance evaluation, and final testing. The selection of appropriate splitting methods significantly impacts on the model accuracy. In this study, a combination of the holdout method and k-fold cross-validation was employed [58]. Fig. 3 describes the methodology used for splitting the dataset in the present work.

Initially, the holdout method is used to randomly divide the training dataset (60 %) from the test dataset (40 %). The test dataset is reserved for the final evaluation of the model's performance, ensuring an unbiased assessment of its generalization ability. Subsequently, the training dataset is further divided using 5-fold cross-validation, where the dataset is divided into 5 equal-sized folds, with each fold serving as a validation set in a rotating manner. The model is trained on different combinations of folds, and the performance is evaluated on the corresponding validation sets. After completing all iterations, the performance metrics obtained from each validation dataset are averaged to provide an overall estimate of the model's performance. K-fold cross-validation helps to provide a more robust evaluation of the model by assessing its performance on different subsets of the data. It helps identifying potential issues like overfitting or data bias and allows for a more accurate estimate of the model's performance metrics by reducing the variance associated with using a single train-test split.

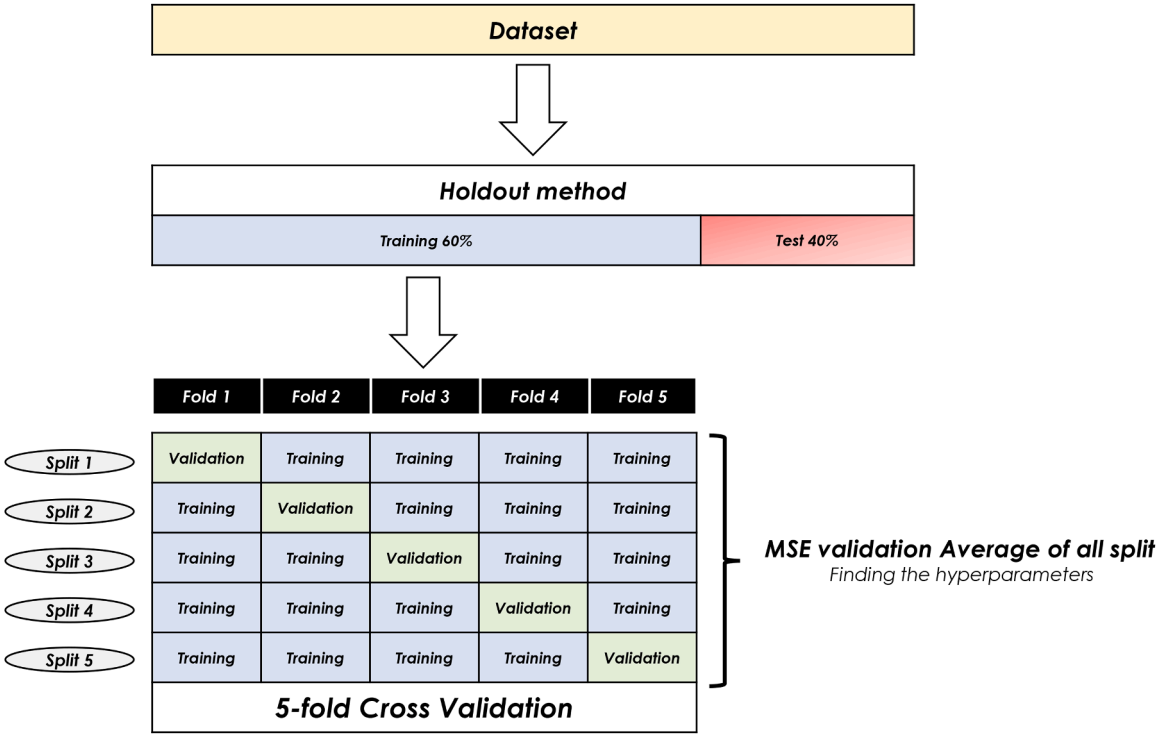


Fig. 3. Integration of holdout and k-fold cross validation method for the dataset division, with the percentage of training and test set.

3.3.2. Neural network hyperparameters selection

As described in previous paragraphs, achieving optimal performance in machine learning models involves tuning multiple hyperparameters, such as activation function, number of hidden layers, training algorithm. Several methods are commonly used for this purpose, including Grid Search, Random Search, and Bayesian Optimization Algorithm (BOA) [61].

Grid Search offers a systematic and exhaustive approach to hyperparameter tuning. It explores a preset grid of hyperparameter values, evaluating the model's performance for each combination. This method guarantees that the best configuration within the specified grid is found. However, a major drawback of Grid Search is its computational cost, especially when dealing with a large number of hyperparameters or a wide range of values [61]. Random Search, instead, randomly samples hyperparameter values from predefined distributions, providing a more flexible exploration of the hyperparameter space with fewer iterations compared to Grid Search. However, there is no guarantee that it will find the optimal configuration because it depends on the randomness of the sampling process [61].

Bayesian Optimization Algorithm (BOA) combines probabilistic models and acquisition functions to explore the hyperparameter space. BOA adapts its search based on past evaluations, making it more efficient and capable of finding good solutions with fewer iterations compared to Grid Search and Random Search [62]. For these reasons the BOA was chosen for this work.

The optimization of neural network hyperparameters plays a fundamental role in obtaining optimal prediction from an ANN model. Fig. 4 provides evidence of the significant impact of the optimization process on the model performance when considering the validation dataset. The results reveal a substantial reduction of approximately 99.78 % in prediction error from the initial parameter combinations to the final stage of optimization. As clearly visible in Fig. 4, most of the effect of the Bayesian optimization approach in terms of MSE minimization is reached after 5 iterations. However, further improvements, although limited, can still be obtained after 180 iterations.

In summary, a combination of the holdout method and the k-fold method for dataset division into training and test datasets was used for this work. K-fold cross-validation technique was integrated with a Bayesian optimizer to determine the optimal combination of neural network model parameters. The dataset was divided with 176 tests used for refining ANN parameters and 117 tests dedicated to evaluating the neural network model's performance. A graphical representation of the described methodology is presented in Fig. 5, while Table 5 provides a comprehensive summary of the parameters considered.

4. Model parameters definition

As explained in the introduction, to control the GCI combustion a properly calibrated injection strategy must be actuated. Multiple injections certainly improve the stability and the controllability of the combustion process, but the ignition stage still remains the key point on the GCI management. Because of the lack of an external igniter (i.e., spark plug), the ignition stage of the air-fuel mixture in GCI engines mainly depends on the combination of physical, thermodynamic, and chemical aspects which can be managed through pilot injections. As mentioned in the introduction, innovative approaches, such as ANN, can increase the performance of the standard injection controllers in unconventional applications such as LTC.

To maximize the accuracy and the coherence with the experimental data of the ANN model, both inputs and output of the FNN must be carefully selected. Several works in literature analyzed the engine and environmental parameters which play a role on ID for GCI combustion [39–41], and the authors defined a pattern of variables which can be used to describe the ignition process of GCI combustion. This section defines the ID calculation methodology for this work and briefly describes the selected inputs and their impact on GCI ignition delay.

4.1. Ignition delay calculation

As stated earlier, in this paper a supervised learning approach was employed to tackle the regression problem of estimating the ID starting from engine parameters and calculated variables available in real-time by the ECU. As reported in Eq. (14), the ID represents the distance between the start of injection (SOI) and the start of combustion (SOC).

$$\Delta\theta_{ID} = \theta_{SOC} - \theta_{SOI} \tag{14}$$

Although the definition of ID is straightforward, various approaches have been proposed for its calculation mainly because of different start of combustion (SOC) definitions [41,63–65]. As mentioned before, an

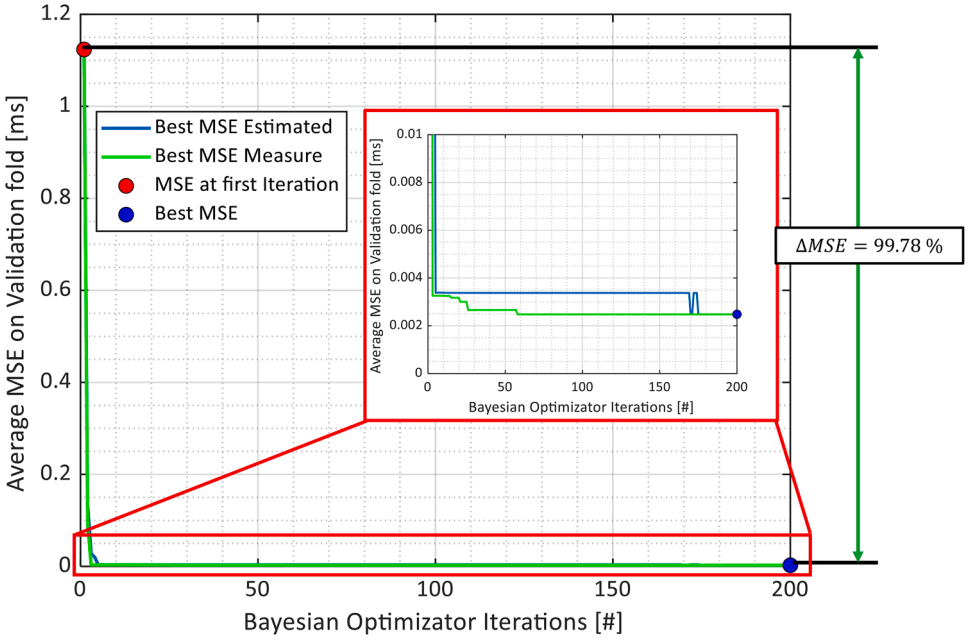


Fig. 4. Variation of Validation MSE during the Bayesian Optimization.

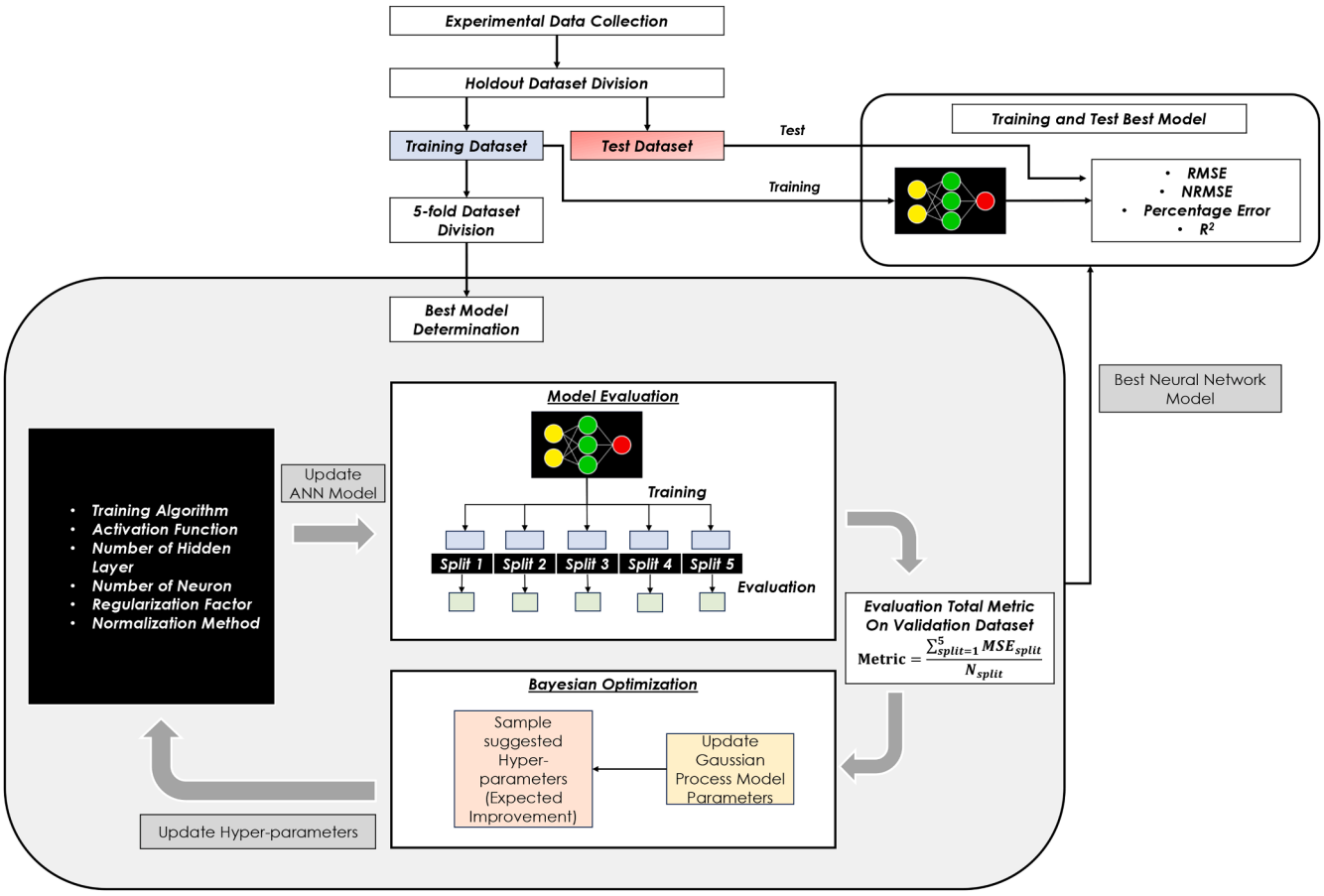


Fig. 5. Graphical Description of the methodology used to find and training the best neural network model.

Table 5
Input for the Bayesian optimization algorithm.

Type of hyperparameter	Values
Activation Functions for each Hidden Layer	sigmoid – tanh -ReLu
Min Number of Hidden Layers	1
Max Number of Hidden Layers	3
Min Number of Neurons for each Hidden Layer	10
Max Number of Neurons for each Hidden Layer	30
Training Algorithms	trainLM – trainSCG - trainBR - trainGDX
Min Regularization Factor value	0
Max Regularization Factor value	1
Input Normalization Methods	zscore – min/max

injection pattern consisting of two pilot injections followed by a main injection was designed to achieve accurate control of GCI combustion. As a result, the ID was defined as the angular distance between the SOI of the first pilot injection $\left(\theta_{SOI_{pilot1}}\right)$ and the *SOC*, which was identified as the crankshaft angle at which the Rate of Heat Release (*ROHR*) reached a threshold of 5 J/deg (θ_{ROHR_5}), Eq. (14). This helps to mitigate the risk of wrong SOC detections due to very small heat released at low temperature and focusing the analysis only on the first combustion stage. A typical GCI *ROHR* trace is reported in Fig. 6 where both the SOC, ID and the SOI of the pilot injections are reported. As it can be seen in the close-up window of Fig. 6, the chosen threshold allows for a repeatable and robust detection of the SOC, thus resulting in a consistent ID estimation for all the considered engine conditions.

$$\Delta\theta_{ID} = \theta_{ROHR\_5} - \theta_{SOI_{pilot1}} \tag{15}$$

Based on the in-cylinder pressure signal and the engine geometry, the *ROHR* calculation was performed following Eq. (16) [66], where V is the combustion chamber volume, p is the pressure inside the cylinder, $dV/d\theta$ and $dp/d\theta$ are their derivatives while γ represents the specific heat ratio. To increase the accuracy of the ROHR calculation, both constant pressure (cp) and constant volume (cv) specific heats were calculated as a function of the bulk gas temperature $\left(T_{cyl}\right)$, following Eq. (17) and Eq. (18) respectively [67].

$$ROHR = \frac{\gamma}{\gamma - 1} p \frac{dV}{d\theta} + \frac{1}{\gamma - 1} V \frac{dp}{d\theta} \tag{16}$$

$$\gamma = \frac{cp\left(T_{cyl}\right)}{cv\left(T_{cyl}\right)} = \frac{cp\left(T_{cyl}\right)}{cp\left(T_{cyl}\right) - R_{air}} \tag{17}$$

$$cp\left(T_{cyl}\right) = 1403.06 - 360.72 * \left(\frac{1000}{T_{cyl}}\right) + 108.24 * \left(\frac{1000}{T_{cyl}}\right)^2 - 10.79 * \left(\frac{1000}{T_{cyl}}\right)^3 \tag{18}$$

4.2. FNN input parameters

Since the aim of this work was to develop and validate a reliable FNN model for ID estimation, a proper input selection was needed. As mentioned before, the more the inputs of the neural network are physically related to the output, the higher the accuracy of the model will be. As a result, the input parameters for the FNN were chosen based on the literature related to the spontaneous ignition of lean air-fuel mixture [37–41].

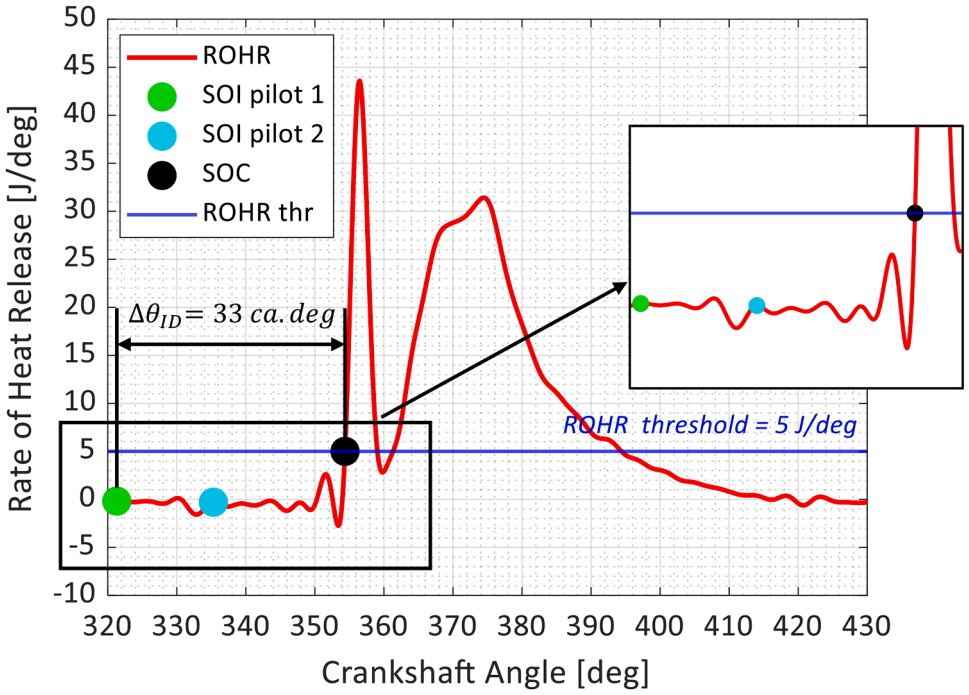


Fig. 6. Experimental calculation of ignition delay in a stable operating point at 12 bar of IMEP and 2000 rpm of engine speed.

Moreover, since the final goal of this work is to achieve a real-time estimation of ID, the input selection for the FNN model was based on parameters that can be directly or indirectly obtained through geometrical considerations and signals from sensors already installed on stock engines, such as crankshaft speed sensor, intake and exhaust manifold pressure and temperature sensors, intake air flowrate sensor, rail pressure sensor, and relative humidity sensor.

Previous work performed by the authors [41] describes the real time calculation of thermodynamic (T_{SOI}) and the chemical-physical (M_{eq}) characteristics of the air-fuel mixture in GCI engines and how these parameters can be used to estimate the ID. According to the definitions given by Silvagni et al. [41] and reported in Eq. (19) and Eq. (20) respectively, such variables summarize the thermodynamic (T_{SOI}) and the chemical-physical (M_{eq}) characteristics of the air-fuel mixture prior to the pilot injection which promotes the ignition in GCI combustion. The thermodynamic parameter T_{SOI} represents the charge temperature in the combustion chamber when the first fuel jet is performed. In Eq. (19), the temperature of the air-fuel mixture when the first fuel jet is introduced into the combustion chamber (SOI_{pilot}) is obtained through the isentropic compression of the charge starting from the intake closure (IVC) temperature. As previously demonstrated by the literature, T_{SOI} can be effectively used to summarize the thermodynamics of the mixture because it already considers aspects with certain impact on ID, such as residuals and fresh air characteristics.

$$T_{SOI} = T_{IVC}\left(\frac{V_{IVC}}{V_{SOI}}\right)^{\gamma-1} \tag{19}$$

Following the definition of chemical-physical variable M_{eq} described by Eq. (20), the chemical potential of the mixture represented by the oxygen content of the fresh air entering the cylinder (M_{O_2}) is decreased by all the chemical components of the air-fuel mixture in the combustion chamber which hinder the combustion propagation, such as the amount of water vapor of the fresh air (M_{H_2O}), the mass of residual gases (M_{RG}), and the amount of external EGR (M_{EGR}). In this formulation of M_{eq}, the model parameters k_{iRG} and k_{EGR} summarize the reduction in the mixture reactivity of the residuals and the external EGR, respectively, because of the internal mixing phenomena. While promising, the high number of experimental conditions needed for model calibration and the decrease in accuracy when running the engine far from the conditions used for calibrating the ID model, represent the limits of the physical-based ID modeling approaches.

$$M_{eq} = M_{O_2} - M_{H_2O} - k_{iRG}M_{RG} - k_{EGR}M_{EGR} \tag{20}$$

As mentioned before, with the aim of maximizing the accuracy of the FNN model and capturing the ID dynamic for GCI combustion, a physical-based approach for input selection was followed. The analysis of ID physical models was used to identify a reasonable set of input variables, reported in Table 6, which describes the autoignition of a lean air-fuel mixture.

Besides the already described chemical and physical parameters, several authors convey that also the pressure of the air-fuel mixture plays a role in the autoignition dynamics. This is also confirmed by Arrhenius-type ID models [41,66], in which the ID of a homogeneous reacting system was described with a function of the in-cylinder pressure, temperature, and equivalence ratio (φ), defined in Eq. (21), by using the air-fuel ratios (AFR) in real and stoichiometric conditions.

$$\varphi = \frac{1}{\frac{AFR}{AFR_{st}}} \tag{21}$$

As a result, one additional parameter, which describes the combustion chamber pressure at the SOI of the first pilot injection P_{SOI}, was added to the input selection. To obtain such variable by using the information of standard sensors mounted on the engine, P_{SOI}, defined in

Table 6
Input for the FNN model.

Variable Name	Relation with the ID
Engine Speed	$\propto^{-1}$
P_{SOI}	$\propto^{-1}$
T_{SOI}	$\propto^{-1}$
P_{rail}	$\propto^{-1}$
m_{O2}	$\propto^{-1}$
m_{EGR}	$\propto$
m_{iRG}	$\propto$
m_{H2O}	$\propto$

Eq. (22), was calculated following an isentropic compression of the charge starting from the pressure at the intake valve closing.

$$P_{SOI} = P_{IVC}\left(\frac{V_{IVC}}{V_{SOI}}\right)^{\gamma} \tag{22}$$

Many studies have documented the impact of the injection pressure on the ID of spontaneously ignited combustions [32–34]. Different injection pressure modifies fuel vaporization, penetration and, consequently, the local stratification of the mixture. As a result, the rail pressure (P_{rail}) should be chosen to maximize the stability and the controllability of the combustion in each engine operating condition. Due to the high sensitivity of ID in GCI combustion to the local air-fuel ratio, to extend the applicability of the presented model, P_{rail} was considered as input for the FNN.

In the field of combustion engine, especially for spontaneous ignition combustion, the flow entering and exiting a combustion chamber is responsible for the fuel stratification and the combustion propagation. As a result, when the combustion chamber geometry and valve timing are fixed, the flow velocity can be considered the main driver which defines the charge motions [35,36]. Therefore, because of the importance of the charge motion in the local air-fuel mixture stratification, the engine speed (which is strictly related to the flow velocity) was selected as additional input for the FNN. Using the set of the input summarized in Table 6, the FNN model was trained and validated using the artificial intelligence tools described in the previous section.

5. Results and discussion

This section presents the FNN model and the cycle-to-cycle ID estimation over the 176 tests used for tuning the parameters of the neural network model and the 117 tests used for model testing. Following the methodology described in Section 3.3, the hyperparameters of the FNN, reported in Table 7, were obtained and optimized using the BOA and k-fold cross-validation.

As can be seen in Fig. 7, the cycle-to-cycle ID estimation using the FNN is in good agreement with the measured ID in both the training (Fig. 7a) and validation (Fig. 7b) datasets for all the engine operating conditions analyzed. As mentioned before, this work mainly focuses on testing the ANN-ID model when running far from the standard engine conditions used for the model calibration. As a result, to obtain a fair assessment of the model accuracy, only stable and steady-state operating conditions were analyzed, neglecting the transient conditions.

Fig. 7a and b also highlight the comparison in terms of root mean square error (RMSE) and percentage normalized RMSE (NRMSE) between the training dataset used for tuning the FNN model and the test dataset respectively. The standard definition of RMSE was applied to the dataset using Eq. (23). Additionally, NMRSE, defined as the ratio of RMSE to the difference between the maximum and minimum values of the actual ID in the dataset, multiplied by 100, was calculated according to Eq. (24).

$$RMSE = \sqrt{\frac{\sum_{p=1}^{N_{sample}}\left(y_p - AptCommand0177_{;p}\right)^2}{N_{sample}}} \tag{23}$$

Table 7
Optimized hyperparameters obtained from the BOA and k-fold cross validation.

Type of Hyperparameter	Values
Activation Functions 1st Hidden Layer	Tanh
Activation Functions 2nd Hidden Layer	sigmoid
Number of Hidden Layers	2
Number of Neurons 1st Hidden Layer	11
Number of Neurons 2nd Hidden Layer	21
Training Algorithms	trainBR
Regularization Factor value	Automatically selected from trainBR
Input Normalization Methods	zscore

$$NRMSE\ [\%] = \left(\frac{RMSE}{\max(y) - \min(y)}\right) * 100 \tag{24}$$

The performance parameters (RMSE and NRMSE) used to evaluate the ID estimation accuracy in both training and test datasets have the same order of magnitude and are comfortably below the self-imposed limits of 0.5 ms and 5 % for RMSE (maximum value of 0.05 ms) and NRMSE (maximum value 2.15 %) respectively. To further confirm the accuracy of the ID estimation realized by using the FNN model, the percentage error (*PE*) calculated following the Eq. (25) on a sample-to-sample basis, was performed.

$$Percentage\ Error_{\ p}\ [\%] = 100 * \left(\frac{y_p - AptCommand0177_{;p}}{y_p}\right) \tag{25}$$

Fig. 8a and b show the *PE* for both the training and test datasets, respectively. As it can be noticed, the average *PE* value is close to zero with a standard deviation below than 1.7 % in for both datasets. It is important to highlight the occurrence of occasional deviations in *PE* of approximately ±10 %. These spikes depend not only on the model but also on the cyclic dispersion present in the target. Despite representing a decrease in accuracy of the ID modeling, this specific behavior can be considered positive because it means that the FNN model approach tends to avoid overfitting.

To further evaluate the accuracy of ID estimation using the machine learning methodology proposed in this work, the R-squared (R^2) was computed for both training and test datasets using the definition outlined in Eq. (22), where the $\bar{y}$ indicates the average of the measured ID. As well-reported in the literature, R^2 is commonly calculated to quantify the variance between the model output and the actual ID.

$$R^2 = 1 - \frac{\sum_{p=1}^{N_{sample}}\left(AptCommand0177_{;p} - y_p\right)^2}{\sum_{p=1}^{N_{sample}}\left(\bar{y} - y_p\right)^2} \tag{22a}$$

Fig. 9a and b show a cycle-to-cycle comparison between the measured and the estimated ID and the R^2 for both training and test datasets. As it can be noticed, the R^2 remains notably high at 0.99 in both datasets, indicating that the model is capable of describing the ID even when an independent (and unseen) test dataset is used. As a consequence, the presented results in terms of accuracy in ID estimation clearly demonstrate the effective predictive prowess of the FNN ID model, attributed to the employed methodology and the thoughtfully selected input features.

When working with machine learning models, it becomes crucial to establish methodologies that provide a deeper understanding of the relationships between input and output variables of the model. One of the most frequently employed methodologies is the evaluation of the relative importance (*RI*) of each input within the model.

Several approaches are reported in the literature to calculate *RI* [45, 68,69]. Among them, the selected *RI* evaluation methodology consists of analyzing the accuracy of the FNN model in terms of RMSE while performing a sensitivity analysis of each input. Such analysis involves introducing a variation of ±10 % for each input, one by one, and then computing the RMSE by comparing the output of the model with the estimated ID obtained using the unaltered input. The individual RMSE is then summed. and the *RI* subsequently calculated following Eq. (23).

$$RI = 100 * \left(\frac{RMSE_i}{\sum_i^{N_{input}} RMSE_i}\right) \tag{23a}$$

Fig. 10 displays the input sensitivity analysis obtained through the previously described methodology for the FNN ID model presented in this work. By looking at Fig. 10, it is evident that the thermodynamic parameters (P_{SOI} and T_{SOI}) and the charge chemical potential (m_{O2}), as well as the engine speed are the most influential variables in the model. Such findings confirm the evidence described in the literature where ID physical models were discussed [41]

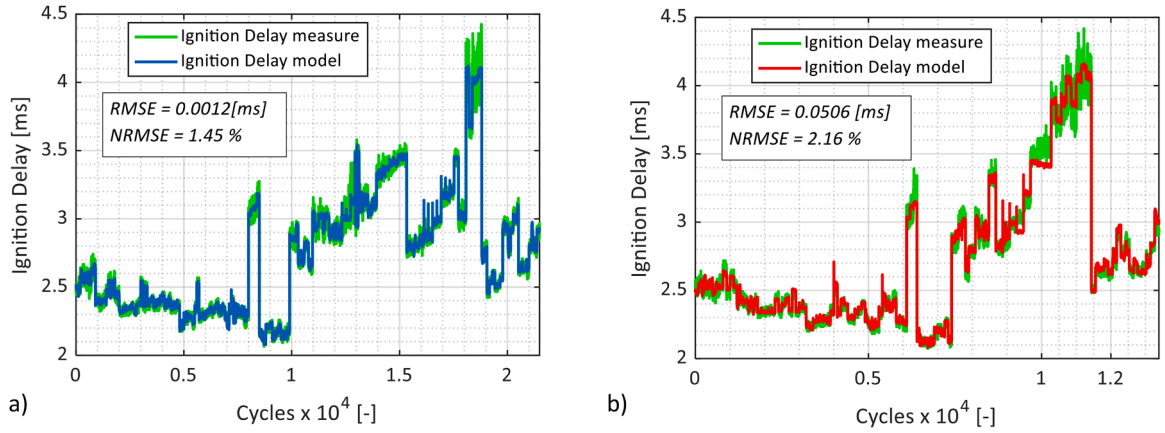


Fig. 7. Comparison of RMSE and percentage NRMSE between training dataset (a) and test dataset (b).

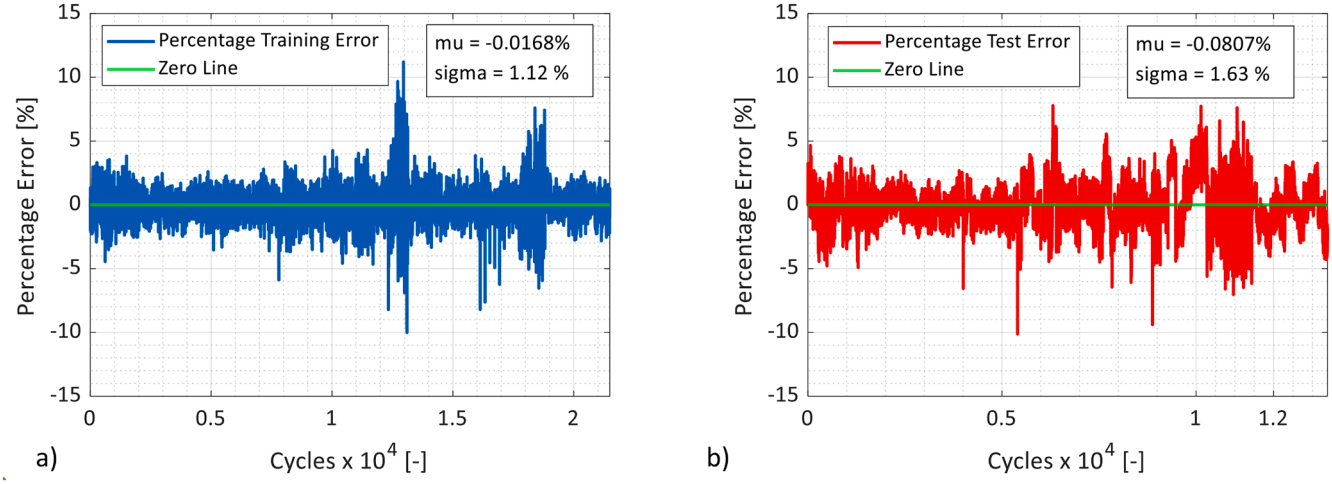


Fig. 8. Comparison of percentage error between training dataset (a) and test dataset (b).

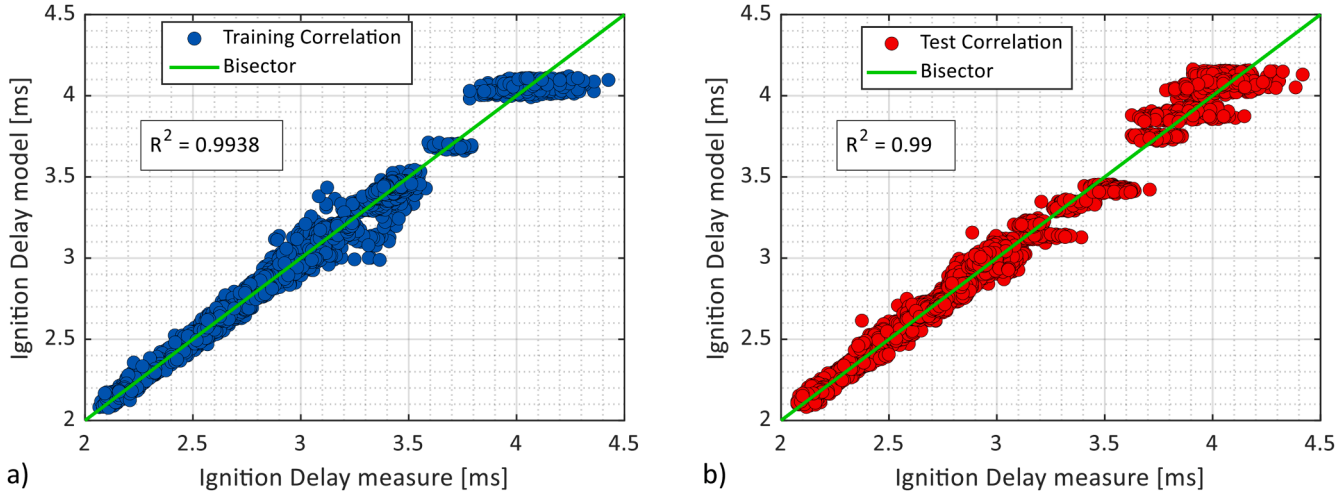


Fig. 9. Comparison of correlation coefficient between training dataset (a) and test dataset (b).

6. Conclusion and future work

In this study, an innovative approach using feedforward neural networks was employed to predict ignition delay in Gasoline Compression Ignition engines. The model selection process involved a meticulous methodology, based on a holdout method to partition the training dataset from the test dataset. A total of 176 engine conditions were allocated for the training phase, while the remaining 117 tests were reserved for the evaluation of the model performance. Starting from the existing physical-based modeling approaches, with the aim of obtaining

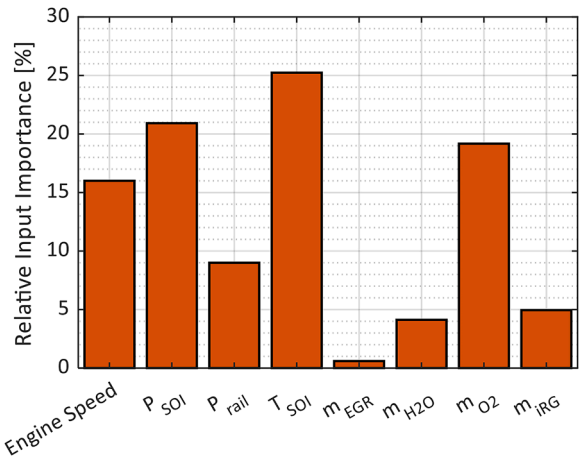


Fig. 10. Importance of each input in the neural network model.

a real-time ID estimation, the input variables of the FNN were identified and calculated as functions of engine parameters and signals typically available in a stock engine layout. To enhance the robustness of the model and identify the most effective hyperparameter combination, an innovative hybrid approach, combining a cross-validation methodology with the Bayesian Optimization Algorithm, was applied.

The analysis of the results demonstrates that the FNN methodology described in this work can replicate the ID of GCI combustion, overcoming the limitations of the existing ID physical models when running far from the standard engine conditions used for the model calibration. The achieved RMSE of 0.05 ms, NRMSE percentage of 2.16 %, and an R^2 of 0.99 on the test dataset for estimating the ID confirm the effectiveness of the proposed approach in describing the GCI ignition dynamics across a wide range of engine and environmental conditions. Furthermore, in alignment with the physical-based approaches for ID estimation, the input sensitivity analysis reveals that the most important variables within the model are the thermodynamic, the engine speed, and the chemical potential of the in-cylinder charge, which further emphasizes the robustness of the model. Furthermore, the input selection for the ANN-model design allows considering this approach suitable for control-oriented applications.

Since one of the main issues with ANN-based models in industrial applications is low repeatability, future work will focus on testing the proposed ID model across different cylinders of the same engine and on various engine platforms. Moreover, a direct comparison between the most accurate physical-based ID models and the proposed ANN-based approach will highlight the benefits and drawbacks of machine learning in the field of advanced combustions.

The integration of the proposed ANN-based model into the injection control strategy will be also performed in the existing open ECU. After integrating the FNN ID estimation model into the existing engine control structure, with the aim of increasing the GCI combustion stability, an innovative injection controller, which combines standard and neural-network modeling approaches, will be developed and tested covering the conventional GCI engine operating range and environmental conditions, as well as ultra-low load and speed engine conditions, where controlling GCI combustion is particularly challenging. The final goal of the activity will be focused on increasing the GCI combustion stability and engine reliability running extreme conditions.

CRediT authorship contribution statement

Rossi Alessandro: Writing – original draft, Visualization, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Silvagni Giacomo:** Writing – original draft, Visualization, Methodology, Investigation, Data curation, Conceptualization. **Ravaglioli Vittorio:** Writing – review & editing, Supervision, Resources, Methodology. **Corti Enrico:** Writing – review & editing, Resources, Project administration.

Declaration of competing interest

all the authors agree that there is no conflicts of interest on this publication.

Data availability

Data will be made available on request.

References

[1] T.v. Johnson, Diesel emission control in review, SAE Int. J. Fuel. Lubr. 1 (2008), https://doi.org/10.4271/2008-01-0069, 2008-01–0069.
[2] Johnson, T.; Joshi, A. Review of vehicle engine efficiency and emissions.; March 28 2017.
[3] DieselNet EU: Cars and Light trucks available online: https://dieselnet.com/standards/eu/ld.php (accessed on 24 October 2022).
[4] J.M. Luján, P. Piqueras, J. de la Morena, F. Redondo, Experimental characterization of real driving cycles in a light-duty diesel engine under different dynamic conditions, Appl. Sci. 12 (2022) 2472, https://doi.org/10.3390/app12052472.
[5] Christoph Böhringer The Kyoto Protocol: A Review and Perspectives 2003, 19, 451–466.
[6] C. Bataille, M. Åhman, K. Neuhoff, L.J. Nilsson, M. Fischedick, S. Lechtenböhmer, B. Solano-Rodriquez, A. Denis-Ryan, S. Stiebert, H. Waisman, et al., A review of technology and policy deep decarbonization pathway options for making energy-intensive industry production consistent with the Paris Agreement, J. Clean. Prod. 187 (2018) 960–973, https://doi.org/10.1016/j.jclepro.2018.03.107.
[7] European Union Regulation (EC). No 443/2009 of the European Parliament and of the Council of 23 April 2009 setting emission performance standards for new passenger cars as part of the community's integrated approach to reduce CO2 emissions from light-duty vehicles. 2009.
[8] Mock P CO2 Emission standards for passenger cars and light-commercial vehicles in the European Union Available online: https://theicct.org/sites/default/files/publications/EU-LCV-CO2-2030_ICCTupdate_20190123.pdf (accessed on 24 October 2022).
[9] European Commision, Communication from the Commision to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Brussels, 2019.
[10] European Union Regulation (EU). 2019/631 of the European Parliament and of the Council of 17 April 2019 setting CO2 emission performance standards for new passenger cars and for New light commercial vehicles, and repealing regulations (EC) No 443/2009 and (EU) No 510/2011 2019.
[11] F. Yan, J. Ling, C. Dai, Y. Wang, J. Li, Research on the adaptability of vehicles with different technical types to the China 6 regulation standard, in: E3S Web of Conferences 268, 2021, p. 01039, https://doi.org/10.1051/e3sconf/202126801039.
[12] J. Pavlovic, B. Ciuffo, G. Fontaras, V. Valverde, A. Marotta, How much difference in type-approval CO_2 emissions from passenger cars in Europe can Be expected from changing to the new test procedure (NEDC vs. WLTP)? Transp. Res. Part A Policy Pract. 111 (2018) 136–147, https://doi.org/10.1016/j.tra.2018.02.002.
[13] Ciuffo, B.; Marotta, A.; Tutuianu, M.; Anagnostopoulos, K.; Fontaras, G.; Pavlovic, J.; Serra, S.; Tsiakmakis, S.; Zacharof, N. The Development of the World-Wide Harmonized Test Procedure For Light Duty Vehicles (WLTP) and the Pathway for Its Implementation Into the EU Legislation.; Washington DC, January 2015.
[14] R. García-Contreras, J.A. Soriano, P. Fernández-Yáñez, L. Sánchez-Rodríguez, C. Mata, A. Gómez, O. Armas, M.D. Cárdenas, Impact of regulated pollutant emissions of euro 6d-temp light-duty diesel vehicles under real driving conditions, J. Clean. Prod. 286 (2021) 124927, https://doi.org/10.1016/j.jclepro.2020.124927.
[15] European Commision *Commission Regulation (EU) 2018/1832 of 5 November 2018 amending Directive 2007/46/EC of the European Parliament and of the Council, Commission Regulation (EC) No 692/2008 and Commission Regulation (EU) 2017/1151 for the purpose of improving the emission type approval tests and procedures for light passenger and commercial vehicles, including those for In-service conformity and real-driving emissions and introducing devices for monitoring the consumption of fuel and electric energy.*
[16] Roberts, P., Mumby, R., Mason, A., Redford-Knight, L. et al., RDE Plus - The Development of a Road, Rig and Engine-in-the-Loop Test Methodology for Real Driving Emissions Compliance, SAE Technical Paper 2019-01-0756, 2019, https://doi.org/10.4271/2019-01-0756.
[17] L. Situ, Electric vehicle development: the past, present & future, in: Proceedings of the 2009 3rd International Conference on Power Electronics Systems and Applications (PESA), 2009, pp. 1–3.
[18] R. Zhang, S. Fujimori, The role of transport electrification in global climate change mitigation scenarios, Environ. Res. Lett. 15 (2020) 034019, https://doi.org/10.1088/1748-9326/ab6658.

[19] P. Piatkowski, W. Puszkiewicz, Electric vehicles: problems or solutions, J. Mech. Energy Eng. 2 (2018) 59–66, https://doi.org/10.30464/jmee.2018.2.1.59.

[20] J.A. Sanguesa, V. Torres-Sanz, P. Garrido, F.J. Martinez, J.M. Marquez-Barja, A review on electric vehicles: technologies and challenges, Smart Citi. 4 (2021) 372–404, https://doi.org/10.3390/smartcities4010022.

[21] acea DRIVING MOBILITY FOR EUROPE fuel types of new cars: battery electric 9.9%, hybrid 22.6% and petrol 38.5% market share in Q2 2022 Available online: https://www.acea.auto/fuel-pc/fuel-types-of-new-cars-battery-electric-9-9-hybrid-22-6-and-petrol-38-5-market-share-in-q2-2022/(accessed on 13 February 2023).

[22] Bureau of Transportation Statistics Hybrid-Electric, Plug-in hybrid-electric and electric vehicle sales Available online: https://www.bts.gov/content/gasoline-hybrid-and-electric-vehicle-sales (accessed on 13 February 2023).

[23] A.K. Agarwal, A.P. Singh, R.K. Maurya, Evolution, challenges and path forward for low temperature combustion engines, Prog. Energy Combust. Sci. 61 (2017) 1–56, https://doi.org/10.1016/j.pecs.2017.02.001.

[24] K. Cho, E. Latimer, M. Lorey, D.J. Cleary, M. Sellnau, Gasoline fuels assessment for Delphi's second generation Gasoline direct-injection compression ignition (GDCI) multi-cylinder engine, SAE Int. J. Eng. 10 (2017), https://doi.org/10.4271/2017-01-0743, 2017-01–0743.

[25] M. Sellnau, J. Sinnamon, K. Hoyer, H. Husted, Gasoline direct injection compression ignition (GDCI) - diesel-like efficiency with low CO2 emissions, SAE Int. J. Eng. 4 (2011), https://doi.org/10.4271/2011-01-1386, 2011-01–1386.

[26] M.C. Sellnau, J. Sinnamon, K. Hoyer, H. Husted, Full-time gasoline direct-injection compression ignition (GDCI) for high efficiency and low NOx and PM, SAE Int. J. Eng. 5 (2012), https://doi.org/10.4271/2012-01-0384, 2012-01–0384.

[27] M. Krishnamoorthi, R. Malayalamurthi, Z. He, S. Kandasamy, A review on low temperature combustion engines: performance, combustion and emission characteristics, Renew.Sustain. Energy Rev. 116 (2019) 109404, https://doi.org/10.1016/j.rser.2019.109404.

[28] A.K. Agarwal, A.P. Singh, A. García, J. Monsalve-Serrano, Challenges and opportunities for application of reactivity-controlled compression ignition combustion in commercially viable transport engines, Prog. Energy Combust. Sci. 93 (2022) 101028, https://doi.org/10.1016/j.pecs.2022.101028.

[29] J.E. Dec, Advanced compression-ignition engines—Understanding the in-cylinder processes, Proc. Combust. Inst. 32 (2009) 2727–2742, https://doi.org/10.1016/j.proci.2008.08.008.

[30] K. Cung, A.A. Moiz, E.M. Smith, D.C. Bitsis, A. Michlberger, T. Briggs, J. Miwa, Gasoline compression ignition (GCI) combustion of pump-grade Gasoline fuel under high compression ratio diesel engine, Transport. Eng. 4 (2021) 100066, https://doi.org/10.1016/j.treng.2021.100066.

[31] H. Wei, J. Hua, M. Pan, D. Feng, L. Zhou, J. Pan, Experimental investigation on knocking combustion characteristics of gasoline compression ignition engine, Energy 143 (2018) 624–633, https://doi.org/10.1016/j.energy.2017.11.020.

[32] N.X. Khoa, Y. Putrasari, D.N. Vu, N.H.X. Duy, O. Lim, The Effect of Control Strategies On the Gasoline Compression Ignition (GCI) Engine: Injection Strategy, Exhaust Residual Gas Strategy, Biodiesel Addition Strategy, and Oxygen Content Strategy, 2022, pp. 27–71.

[33] A.K. Agarwal, V.S. Solanki, M. Krishnamoorthi, Gasoline compression ignition (GCI) combustion in a light-duty engine using double injection strategy, Appl. Therm. Eng. 223 (2023) 120006, https://doi.org/10.1016/j.applthermaleng.2023.120006.

[34] Stola, F., Ravaglioli, V., Silvagni, G., Ponti, F. et al., Analysis of the Effects of Injection Pressure Variation in Gasoline Partially Premixed Combustion, SAE Technical Paper 2021-01-0517, 2021, https://doi.org/10.4271/2021-01-0517.

[35] Li, C., Yin, L., Shamun, S., Tuner, M. et al., "Transition from HCCI to PPC: the Sensitivity of Combustion Phasing to the Intake Temperature and the Injection Timing with and without EGR," SAE Technical Paper 2016-01-0767, 2016, https://doi.org/10.4271/2016-01-0767.

[36] G. Belgiorno, N. Dimitrakopoulos, G. di Blasio, C. Beatrice, P. Tunestål, M. Tunér, Effect of the engine calibration parameters on gasoline partially premixed combustion performance and emissions compared to conventional diesel combustion in a light-duty euro 6 engine, Appl. Energy 228 (2018) 2221–2234, https://doi.org/10.1016/j.apenergy.2018.07.098.

[37] Stola, F., Ravaglioli, V., Silvagni, G., Ponti, F. et al., Injection Pattern Investigation for Gasoline Partially Premixed Combustion Analysis, SAE Technical Paper 2019-24-0112, 2019, https://doi.org/10.4271/2019-24-0112.

[38] M. Alabbad, Y. Li, K. AlJohani, G. Kenny, K. Hakimov, M. Al-lehaibi, A.-H. Emwas, P. Meier, J. Badra, H. Curran, et al., Ignition delay time measurements of diesel and gasoline blends, Combust. Flame 222 (2020) 460–475, https://doi.org/10.1016/j.combustflame.2020.09.008.

[39] J.J. Hernández, J. Sanz-Argent, J.M. Carot, J.M. Jabaloyes, Ignition delay time correlations for a diesel fuel with application to engine combustion modelling, Int. J. Eng. Res. 11 (2010) 199–206, https://doi.org/10.1243/14680874JER06209.

[40] Y. Lee, S. Lee, K. Min. Ignition Delay Model of Multiple Injections in, CI Engines," SAE Technical Paper, 2019-24-0071, 2019. https://doi.org/10.4271/2019-24-0071.

[41] G. Silvagni, V. Ravaglioli, S. Falfari, F. Ponti, V. Mariani, Development of a control-oriented ignition delay model for GCI combustion, Energ. (Basel) 15 (2022) 6470, https://doi.org/10.3390/en15176470.

[42] I. Veza, A. Afzal, M.A. Mujtaba, A. Tuan Hoang, D. Balasubramanian, M. Sekar, I. M.R. Fattah, M.E.M. Soudagar, A.I. EL-Seesy, D.W. Djamari, et al., Review of artificial neural networks for gasoline, diesel and homogeneous charge compression ignition engine, Alexandr. Eng. J. 61 (2022) 8363–8391, https://doi.org/10.1016/j.aej.2022.01.072.

[43] Q. Huang, J. Liu, C. Ulishney, C.E. Dumitrescu, On the use of artificial neural networks to model the performance and emissions of a heavy-duty natural gas spark ignition engine, Int. J. Eng. Res. 23 (11) (2022) 1879–1898, https://doi.org/10.1177/14680874211034409.

[44] A.N. Bhatt, N. Shrivastava, Application of artificial neural network for internal combustion engines: a State of the art review, Arch. Comput. Method. En. 29 (2022) 897–919, https://doi.org/10.1007/s11831-021-09596-5.

[45] J. Liu, H. Wang, Machine learning assisted modeling of mixing timescale for LES/PDF of high-Karlovitz turbulent premixed combustion, Combust. Flame 238 (2022) 111895, https://doi.org/10.1016/j.combustflame.2021.111895.

[46] W. Wang, P. Jones, D. Partridge, Assessing the impact of input features in a feedforward neural network, Neur. Comput. Appl. 9 (2000) 101–112, https://doi.org/10.1007/PL00009895.

[47] U. Anders, O. Korn, Model selection in neural networks, Neur. Netw. 12 (1999) 309–323, https://doi.org/10.1016/S0893-6080(98)00117-8.

[48] J. Awwalu, A comparison between Neural network and Support Vector machine, Int. J. Trend Rese. Develop. 6 (2019), https://doi.org/10.13140/RG.2.2.13811.96809.

[49] P. Refaeilzadeh, L. Tang, H. Liu. Cross-validation, Encyclopedia of Database Systems, Springer US, Boston, MA, 2009, pp. 532–538. https://doi.org/10.1007/978-0-387-39940-9_565.

[50] H. Wang, C. Ji, T. Su, C. Shi, Y. Ge, J. Yang, S. Wang, Comparison and implementation of machine learning models for predicting the combustion phases of hydrogen-enriched Wankel rotary engines, Fuel 310 (2022) 122371, https://doi.org/10.1016/j.fuel.2021.122371.

[51] Julianna D. Supervised vs. Unsupervised learning: what's the difference? Available online: https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning (accessed on 25 February 2023).

[52] R. Sathya, Annamma Abraham comparison of supervised and unsupervised learning algorithms for pattern classification, Int. J. Adv. Res. Artifi. Intell. 2 (2013) 34–38, https://doi.org/10.14569/IJARAI.2013.020206.

[53] Z. Zhang, F. Feng, T. Huang, FNNS: an effective feedforward neural network scheme with random weights for processing large-scale datasets, Appl. Sci. 12 (2022) 12478, https://doi.org/10.3390/app122312478.

[54] B. Ding, H. Qian, J. Zhou, Activation functions and their characteristics in deep neural networks, in: Proceedings of the 2018 Chinese Control And Decision Conference (CCDC), IEEE, 2018, pp. 1836–1841, https://doi.org/10.1109/CCDC.2018.8407425.

[55] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neur. Netw. 4 (1991) 251–257, https://doi.org/10.1016/0893-6080(91)90009-T.

[56] Ajitesh Kumar Difference between online & batch learning Available online: https://vitalflux.com/difference-between-online-batch-learning/(accessed on 3 June 2023).

[57] Manish Chablani Gradient descent algorithms and adaptive learning rate adjustment methods Available online: https://towardsdatascience.com/gradient-descent-algorithms-and-adaptive-learning-rate-adjustment-methods-79c701b086be (accessed on 3 June 2023).

[58] H. Liu, On the Levenberg-Marquardt training method for feed-forward neural networks, in: ComputationProceedings of the 2010 Sixth International Conference on Natural Computation, IEEE, 2010, pp. 456–460, https://doi.org/10.1109/ICNC.2010.5583151.

[59] C.K. Arthur, V.A. Temeng, Y.Y. Ziggah, Performance evaluation of training algorithms in backpropagation neural network approach to blast-induced ground vibration prediction, Ghana Mining J. 20 (2020) 20–33, https://doi.org/10.4314/gm.v20i1.3.

[60] S.C. Nayak, B.B. Misra, H.S. Behera, impact of data normalization on stock index forecasting, Int. J. Comput. Inform. Syst. Ind. Manage. Applic. 6 (2014) 257–269.

[61] H. Wang, C. Ji, T. Su, C. Shi, Y. Ge, J. Yang, S. Wang, Comparison and implementation of machine learning models for predicting the combustion phases of hydrogen-enriched Wankel rotary engines, Fuel 310 (2022) 122371, https://doi.org/10.1016/j.fuel.2021.122371.

[62] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. de Freitas, Taking the Human out of the loop: a review of bayesian optimization, Proceed. IEEE 104 (2016) 148–175, https://doi.org/10.1109/JPROC.2015.2494218.

[63] P. Fabrizio, R. Vittorio, C. Nicolò, D.C. Matteo, Diesel engine combustion sensing methodology based on vibration analysis, J. Eng. Gas Turb. Power (2014) 136, https://doi.org/10.1115/1.4027363.

[64] S. Oh, K. Min, M. Sunwoo, Real-time start of a combustion detection algorithm using initial heat release for direct injection diesel engines, Appl. Therm. Eng. 89 (2015) 332–345, https://doi.org/10.1016/j.applthermaleng.2015.05.079.

[65] J.M. Luján, V. Bermúdez, C. Guardiola, A. Abbad, A methodology for combustion detection in diesel engines through In-cylinder pressure derivative signal, Mech. Syst. Signal. Process. 24 (2010) 2261–2275, https://doi.org/10.1016/j.ymssp.2009.12.012.

[66] J. Heywood, Internal Combustion Engine Fundamentals, Ed., McGraw-Hill Education, New York, 1998.

[67] Pipitone, E., Beccari, A., and Beccari, S., "Reliable TDC position determination: a comparison of different thermodynamic methods through experimental data and simulations," SAE Technical Paper 2008-36-0059, 2008, https://doi.org/10.4271/2008-36-0059.

[68] J. OLDEN, An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data, Ecol. Modell. (2004), https://doi.org/10.1016/S0304-3800(04)00156-5.

[69] Hassan Ibrahim; Omar M. Ibrahim Comparison of Methods for Assessing the Relative Importance of Input Variables in Artificial Neural Networks. *The Journal of Applied Sciences Research* 2013, *9*, 5692–5700.