Neural-Network Based Approach for Real-Time Control of BMEP and MFB50 in a Euro 6 Diesel Engine

2017-24-0068

Neural-Network Based Approach for Real-Time Control of BMEP and MFB50 in a Euro 6 Diesel Engine

Roberto Finesso, Ezio Spessa, and Yixin Yang
Politecnico di Torino

Giuseppe Conte and Gennaro Merlino
General Motors Global Propulsion Systems

Abstract

A real-time approach has been developed and assessed to control BMEP (brake mean effective pressure) and MFB50 (crank angle at which 50% of fuel mass has burnt) in a Euro 6 1.6L GM diesel engine. The approach is based on the use of feed-forward ANNs (artificial neural networks), which have been trained using virtual tests simulated by a previously developed low-throughput physical engine model. The latter is capable of predicting the heat release and the in-cylinder pressure, as well as the related metrics (MFB50, IMEP – indicated mean effective pressure) on the basis of an improved version of the accumulated fuel mass approach. BMEP is obtained from IMEP taking into account friction losses. The low-throughput physical model does not require high calibration effort and is also suitable for control-oriented applications. However, control tasks characterized by stricter demands in terms of computational time may require a modeling approach characterized by a further lower throughput. To this aim, feed-forward NNs have been trained to predict MFB50 and BMEP using a large dataset of virtual tests generated by the well-calibrated low-throughput physical engine model. The real-time approach has also been applied to derive the start of injection of the main pulse and the injected fuel quantity to achieve specific targets of MFB50 and BMEP. The accuracy of the real-time approach has been assessed based on experimental data taken at GM-GPS (General Motors – Global Propulsion Systems) facilities and its computational time has been compared to that of the low-throughput physical engine model, at steady-state and transient conditions over the WLTP cycle.

Introduction

Interest in developing model-based algorithms for combustion control in IC engines is growing more and more among car manufacturers, due to the increasing computational performance of modern ECUs (Engine Control Units). The adoption of a model-based control approach, instead of a traditional map-based one, can in fact lead to a reduction in the experimental effort required for map calibration, with a consequent decrease in the related costs. This is especially true for a diesel engine, which is characterized by a large number of control variables (e.g., injection timings, injection quantities, rail pressure, variable geometry turbocharger (VGT) rack position, exhaust gas recirculation (EGR) rate, swirl ratio, etc.).

Therefore, the development of control-oriented models for the prediction of combustion parameters and, more in general, of the performance of a diesel engine has become of great interest in the last few years.

In general, engine simulation can be carried out with different degrees of detail. For example, the simulation of the fluid-dynamics in the pipes, manifolds and in the combustion chamber, as well as of the combustion and emission formation processes, can be carried out using multidimensional, one-dimensional or zero-dimensional approaches.

3D-CFD (Computer Fluid-Dynamics) methods have the potential of reproducing the physical and chemical processes that take place in the engine pipes and in the chamber during the injection-combustion process, but they require a high computational time. Thus, these methods are generally used to simulate specific engine sub-systems (see [1-9]).

Instead, the complete simulation of the engine system is usually carried out by coupling 1D-CFD models (which are adopted for the simulation of the flow inside the pipes) with zero-dimensional models that simulate the cylinders, injectors, valves, and turbocharger performance. Some examples can be found in the literature [10-14]. However, this approach requires a computational time that is not currently suitable to develop a model-based control.

Mean-value engine models [15, 16] simulate the intake/exhaust engine variables by means of steady-state and empirical correlations, and are capable of simulating the combustion and emission formation processes with a good level of detail [16]. These methods offer the opportunity of further decreasing the computational time compared to 1D-CFD methods, while guaranteeing at the same time a good predictive capability at steady-state and mildly transient engine operating conditions. Moreover, they are generally physically consistent, so that they do not require a high calibration effort, and at the same time their accuracy is still acceptable outside the calibration range [16]. Therefore, they can be considered as good candidates for the development of model-based control algorithms.

Finally, a last category of models, which are often used in the field of engine design and control, includes artificial intelligence systems such as support vector machines (SVM) and artificial neural networks (ANNs). These methods, which can also be referred to as

"black box" or "gray box" approaches, do not require the detailed physical knowledge of the investigated process and are able to capture complex nonlinear system behavior with relatively simple mathematical operation. Moreover, they are characterized by a very small computational time, so that they are the best candidates for the development of model-based control algorithms to be implemented in ECUs (Engine Control Units). However, their training usually requires a high number of experimental tests, and their performance is usually not reliable outside the calibration range. Artificial intelligence systems also include methods such as genetic algorithms (GA), which are typically adopted to solve optimization problems.

A lot of applications of this type can be found in the literature. For example, SVMs have been used to reconstruct the indicated torque from crankshaft velocity in a six-cylinder spark ignition engine [17]. GAs have been used to optimize diesel engine emissions and fuel economy with the existing techniques, such as exhaust gas recirculation and multiple injections [18]. Differential evolution algorithms (DEAs) were used to improve the gas exchange process of a single-cylinder compression ignition naturally aspirated engine in [19].

Among these methods, ANNs are one of the more powerful candidates due to their efficiency, flexibility and robustness.

Several applications, based on ANNs, can be found in the literature for engine control and diagnostics. For example, in [20] an explicit back propagation neural network was developed to identify diesel combustion misfire according to the general engine operating parameters. In [21] a combined neural network and fuzzy logic-based control scheme was designed for spark advance control to get maximum brake torque timing. An application based on ANNs for the optimization of the gear shifting strategy can be found in [22].

ANNs have also been employed for air mass flow estimation and the related control of the air-to-fuel ratio [23-25]. In [26], ANNs have been employed as virtual sensors for the tuning and optimization of hydrogen powered vehicle.

Therefore, ANNs have been shown to play an important role for engine simulation and performance prediction. The related research can be found in [27-34]. In [29] a neural network-based methodology to predict HCCI combustion behavior during transient operation was reported. In [30] ANNs were used to predict the brake specific fuel consumption, effective power and exhaust temperature of a gasoline engine. In [33] an ANN-based model was implemented and used to predict combustion product composition and gas thermodynamic properties.

Contribution of the present study

From the previous background, it should be noted that mean-value models and artificial intelligence-based methods, especially ANNs, are the best candidates for the development of model-based control algorithms. The former methods are more physically consistent, so that they do not require a high calibration effort, and are generally characterized by low computational time. The latter approaches require a much shorter computational effort, so that they are more suitable for implementation in an engine ECU. However, their training usually requires a high number of experimental data from a test bench (especially when the number of neurons or the number of layers is high), which would be quite expensive and time consuming. In addition, they are not usually very robust outside the calibration range.

These considerations have suggested the possibility of developing an ANNs-based methodology in which the neural networks are not trained on the basis of experimental data, but using a large dataset of engine conditions simulated by a mean-value engine model, which is therefore used as a virtual engine.

The resulting ANN-based model, therefore, has the aim of replicating the outcomes of the mean-value engine model, but with the advantage of requiring a much shorter computational time, in view of the onboard implementation for model-based control tasks. This approach has therefore been referred to as NNRT (Neural Network-based Real-Time) model.

In this paper, a NNRT model has been developed to predict BMEP (brake mean effective pressure) and MFB50 (crank angle at which 50% of fuel mass has burnt) in a Euro 6 1.6L GM diesel engine. These metrics are in fact widely adopted for combustion control applications in diesel engines.

A real-time mean-value engine model, which was previously developed by the authors in [16], was used for the training of the NNRT model. The mean-value model includes the simulation of chemical heat release and in-cylinder pressure, as well as of the related metrics, such as MFB50, PFP (Peak Firing Pressure), IMEP (Indicated Mean Effective Pressure) and BMEP. This model will be also referred to as "physical engine model" throughout the paper, given its physical consistency.

In the mean-value model, the prediction of the chemical heat release is based on an improved version [16, 35] of the accumulated fuel mass approach, previously presented in [36-40]. This approach assumes that, at any time instant, the rate of chemical energy released by the fuel is proportional to the energy associated with the in-cylinder accumulated fuel mass. The original approach has been physically enhanced in [35] in order to account for the noticeable turbulence effect when high injection pressure levels are adopted at high load conditions. The simulation of the in-cylinder pressure is based on the inversion of a single-zone model [41]. It was shown in [16] that the mean-value engine model requires a very low number of experimental data for training, and is characterized by a computational time that is of the order of a few milliseconds when it is run on a PC. Therefore, it is capable of simulating a very large dataset of engine conditions for the NNRT model training, with a reasonable computational effort.

The NNRT model developed in this paper exploits single hidden layer feed-forward neural networks. The model has the same inputs as the physical one (i.e., fuel injection parameters, thermodynamic conditions in the intake and exhaust manifold, engine speed, etc.) and predicts MFB50 and BMEP as performance indices of combustion phasing and engine load, respectively. Both physical and fast-running models have also been inverted in order to estimate the required SOI_{main} (start of injection of the main pulse) and injected fuel quantity $q_{f,inj}$ to achieve specific targets of MFB50 and BMEP. Model inversion is in fact mandatory for the development of a model-based control algorithm.

Two different DoE (Design of Experiment) approaches, one based on a global variation of the all the engine parameters and the other one based on a local variation around selected engine operating conditions, have been proposed and compared for the training of the

NNRT model. Moreover, a sensitivity analysis has been carried out in order to investigate the influence of the number of neurons on the accuracy of the NNRT model.

The performance of both the physical and NNRT models has been compared at steady state condition and transient conditions, over the WLTP cycle.

Engine setup and experimental activity

The experimental tests for the calibration and validation of the models were conducted on a 1.6L Euro 6 GM diesel engine. The main engine technical specifications are reported in Table 1. The engine was fueled with a diesel oil characterized by an average density of 835 kg/m^3 at 14°C, an average viscosity of 2 mm^2/s at 40°C and an average cetane number equal to 43.

Table 1. Main engine specifications.

Engine type	Euro 6 diesel engine
Displacement	1598 cm^3
Bore x stroke	79.7 mm x 80.1 mm
Rod length	135 mm
Compression ratio	16.0
Valves per cylinder	4
Turbocharger	VGT type
Fuel injection system	Common Rail
Specific power and torque	71 kW/L – 205 Nm/L

The engine is equipped with a short-route cooled EGR system, in which the EGR valve is located upstream from the cooler. A throttle valve is installed upstream from the intake manifold and EGR junction, in order to allow high EGR rates to be obtained when the pressure drop between the exhaust and intake manifolds is not sufficient. Moreover, the EGR circuit is equipped with an EGR cooler bypass, in order to prevent EGR gases from flowing across the cooler under certain driving conditions, e.g., during cold start phases.

The test engine was instrumented with piezoresistive pressure transducers and thermocouples to measure the pressure and temperature at different locations, such as upstream and downstream from the compressor, turbine and intercooler, and in the intake manifold. Thermocouples were also used to measure the temperatures in each exhaust runner. Piezoelectric transducers were installed to measure the pressure time-histories in the combustion chamber of the cylinders.

The experimental tests were carried out on a dynamic test bench at GM-GPS (General Motors – Global Propulsion Systems), in the frame of a research project between the Politecnico di Torino and GM-GPS pertaining to the assessment of control-oriented heat release predictive models [35]. To this aim, several tests were conducted, including:

- Full-Factorial variation tests of p_{int} (intake manifold pressure), SOI_{main} (start of injection of the main pulse), O_2 (intake oxygen concentration) and p_f (injection pressure) at several representative key-points of the NEDC (New European Driving Cycle).
- A full engine map with baseline operating parameters.

Table 2 reports the experimental test matrix used for the calibration of the physical model, and the minimum and maximum values of the main engine parameters for the different operating points.

Table 2. Experimental test matrix used for physical model calibration.

Operating point	Tests		p_f	p_{int}	T_{int}	O_2	λ	X_r	q_{pil2}	q_{pil1}	$q_{f,inj}$	SOI_{pil2}	SOI_{pil1}	SOI_{main}
			bar	bar	°C	%	-	-	mm^3	mm^3	mm^3	CAD	CAD	CAD
1500x2	100	max	500	1.26	78	20.6	4.78	50.1	1.3	1.6	9.18	332.7	342.6	353.2
		min	250	0.95	41	14.7	1.65	1.0	1.3	1.5	7.74	339.4	348.5	358.2
1500x5	100	max	680	1.45	57	20.7	2.81	36.5	1.3	1.6	17.76	331.5	340.7	352.9
		min	380	0.95	42	14.4	1.14	0.5	1.3	1.6	15.42	339.5	348.5	358.2
1500x8	100	max	800	1.68	48	20.8	2.18	17.9	1.3	1.6	28.13	331.4	340.0	355.5
		min	400	1.10	41	18.1	1.11	0.3	1.3	1.6	23.36	337.9	346.4	360.0
2000x2	100	max	580	1.58	90	20.7	5.13	48.2	1.3	1.5	10.11	327.4	339.8	354.7
		min	250	0.85	41	15.3	1.63	0.7	1.3	1.5	7.90	332.8	343.8	356.6
2000x5	100	max	680	1.87	76	20.7	3.48	30.5	1.3	1.5	17.05	326.8	338.3	354.8
		min	380	0.97	50	16.4	1.30	0.5	1.3	1.5	15.02	329.8	340.7	355.4
2000x8	100	max	1100	2.09	67	20.8	2.64	21.0	1.3	1.5	26.64	325.1	335.9	356.1
		min	700	1.34	51	17.2	1.17	0.4	1.3	1.5	23.44	326.8	337.2	356.4
2000x12	100	max	1040	2.46	52	20.8	2.18	0.5	1.3	1.8	36.05	326.2	338.7	351.5
		min	640	1.84	47	20.8	1.75	0.3	1.3	1.8	33.50	327.3	339.3	351.5
Engine map	285	max	1936	2.73	92	20.8	5.77	50.6	1.3	3.1	74.03	311.5	306.3	337.5
		min	305	1.01	34	16.8	1.08	0.1	0.0	1.3	5.47	360.0	353.4	360.2

The key-points are identified by the speed and BMEP (rpm x bar) values. X_r is the EGR rate, defined as the ratio between the inducted EGR mass flow rate and the total inducted mass flow rate; λ is the relative air-to-fuel ratio, q_{pil2} and q_{pil1} indicate the injected fuel quantity of the pilot2 and pilot1 injections, respectively, while $q_{f,inj}$ denotes the total injected fuel quantity. All injected quantities are expressed in mm^3/cyc.

Figure 1 reports the engine map tests, in terms of normalized BMEP and engine speed.

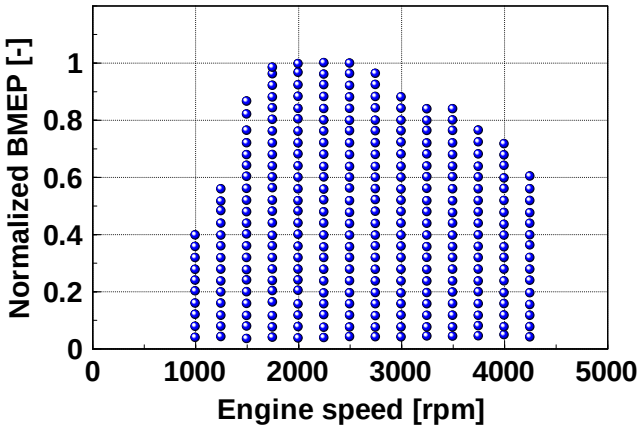


Figure 1. Engine map tests used for the calibration of the physical model.

Physical and NNRT models

This section reports the description of the physical engine model and of the newly proposed NNRT model. The physical engine model was developed by the authors in [16]. However, a summary is reported in the next section for the sake of clarity. It is worth recalling that the physical engine model has been used as a virtual engine in order to simulate the data for the training of the NNRT model. The latter model, therefore, has the aim of replicating the outcomes of the physical engine model, but with the advantage of requiring a much

shorter computational time, in view of the onboard implementation for model-based control tasks.

Physical engine model

The physical engine model that has been used in this study for the training of the NNRT model includes the simulation of:

1. Chemical energy release: the approach is based on an enhanced version [35] of the model previously presented by the authors, which was based on the accumulated fuel mass approach [40]. The input data of the model are the injection parameters, as well as the main thermodynamic conditions in the intake manifold and the engine operating parameters.

2. In-cylinder pressure: the approach is based on the inversion of a single-zone heat release model which requires the net energy release as input; the latter is derived starting from the predicted chemical energy release and estimating the heat transfer between the charge and the walls. Polytropic evolutions are assumed during the compression and expansion phases. The simulation of the in-cylinder pressure allows several parameters, such as PFP and IMEP, to be evaluated.

3. Friction losses: the Chenn-Flynn approach has been used to predict FMEP on the basis of the engine speed and PFP; the simulation of FMEP allows BMEP to be evaluated starting from IMEP.

The detailed description of the model is described in [16]. However, a summary is reported hereafter for the sake of clarity.

Estimation of the chemical energy release Q_{ch}

The chemical energy release has been simulated on the basis of an enhanced version [35] of the baseline model presented by the authors in [40], which was based on the accumulated fuel mass approach.

In particular, the chemical energy release rate for pilot injections is evaluated as follows:

$$\frac{dQ_{ch,pil,j}}{dt}(t) = K_{pil,j}\left[Q_{fuel,pil,j}(t-\tau_{pil,j}) - Q_{ch,pil,j}(t)\right] \quad (1)$$

where $K_{pil,j}$ and $\tau_{pil,j}$ are model calibration quantities related to the combustion rate and to the ignition delay, respectively, and $Q_{fuel,pil,j}$ is the chemical energy associated with the injected fuel mass.

The chemical energy release of the main pulse has instead been simulated as follows:

$$\frac{dQ_{ch,main}}{dt}(t) = K_{1,main}\left[Q_{fuel,main}(t-\tau_{main}) - Q_{ch,main}(t)\right] + K_{2,main}\frac{dQ_{fuel,main}(t-\tau_{main})}{dt} \quad (2)$$

The formulation proposed in Eq. (2) needs an additional calibration parameter with respect to the baseline approach of Eq. (1) (i.e., $K_{2,main}$).

For each injection pulse j, the chemical energy Q_{fuel} associated with the injected fuel quantity is defined as follows:

$$Q_{fuel,j}(t) = \int_{t_{SOI,j}}^{t} \dot{m}_{f,inj}(t) H_L dt \qquad t \le t_{EOI,j} \quad (3)$$

$$Q_{fuel,j}(t) = \int_{t_{SOI,j}}^{t_{EOI,j}} \dot{m}_{f,inj}(t) H_L dt \qquad t > t_{EOI,j} \quad (4)$$

where t_{SOI} is the start of the injection time, t_{EOI} the end of the injection time, H_L the lower heating value of the fuel and $\dot{m}_{f,inj}$ the fuel mass injection rate.

The total chemical energy release is given by the sum of the contributions of all the injection pulses:

$$Q_{ch} = \sum_{j=1}^{n} Q_{ch,j} \quad (5)$$

The model was assessed for the steady-state conditions reported in Table 2. In particular, the optimal τ and K parameters were identified by comparing the predicted and experimentally-derived heat release profiles, and minimizing the sum of errors and the MFB50 prediction error by means of a genetic algorithm (see [35]). The adopted correlation variables for ignition delay were chosen in accordance with the study proposed in [42]. In particular, at the beginning, all the engine variables were included in the correlations, and a sensitivity analysis was carried out in order to exclude the least influential ones, thus a stepwise regression was adopted. This approach was also adopted to identify the engine variables that had to be included in all the correlations reported hereafter.

The following correlations were identified in [35] as a function of the in-chamber thermodynamic quantities at SOI/SOC (start of injection/start of combustion) and other engine variables:

$$K_{pil}\left[\frac{1}{s}\right] = 1.109\, exp\left(\frac{-10254}{T_{SOC,pil}}\right) O_2^{3.80} n^{0.724} q_{pil,tot}^{-0.866} q_{f,inj}^{1.207} \quad (6)$$

$$K_{1,main}\left[\frac{1}{s}\right] = 19.91\; p_f^{0.37} \rho_{SOC,main}^{0.575} n^{0.225} q_{f,inj}^{-0.69} \quad (7)$$

$$K_{2,main}\left[\frac{s}{deg}\right] = 0.296\; p_f^{0.84} \rho_{SOC,main}^{0.59} n^{-0.507} K_{1,main}^{-2.12} \quad (8)$$

$$\tau_{pil}\left[deg\right] = 257\, \rho_{SOI,pil}^{-0.757}\, exp\left(\frac{1559}{T_{SOI,pil}}\right) O_2^{-1.17} \quad (9)$$

$$\tau_{main}\left[deg\right] = 508\; p_f^{-0.648} \rho_{SOI,main}^{-0.95} O_2^{-1.265} n^{0.827} q_{f,inj}^{-0.323} \quad (10)$$

In equations (6-10), ρ_{SOI}, T_{SOI}, ρ_{SOC} and T_{SOC} indicate the in-chamber densities and temperatures evaluated at the start of injection or combustion, respectively, and are expressed in kg/m^3 and K. The injection pressure p_f is expressed in bar, the engine speed n in rpm,

the total injected fuel quantity $q_{f,inj}$ (used as a load parameter) in mm^3/cyc/cyl, the total injected fuel quantity of the pilot shots $q_{pil,tot}$ in mm^3/cyc/cyl and finally the intake oxygen concentration O_2 in %. For the combustion rate parameters, the in-chamber thermodynamic conditions evaluated at SOC were selected as being more representative than those evaluated at SOI. It should be noted that it is necessary to convert the ignition delay values in 's' before their utilization in Eqs. (1-2).

Estimation of the net energy release Q_{net}

The simulation of the in-cylinder pressure during the combustion phase requires the net energy release Q_{net} as input. The latter has been estimated starting from the chemical release, taking into account the heat transfer and fuel evaporation heat effects, according to the following formulation [40]:

$$Q_{net,ht} \cong Q_{ch} \frac{m_{f,inj} H_L - Q_{ht,glob}}{m_{f,inj} H_L} \qquad (11)$$

$$Q_{net} \cong Q_{net,ht} - Q_{f,evap} \qquad (12)$$

where $Q_{net,ht}$ is the result of the scaling of the chemical energy release according to the global heat exchanged by the charge with the walls. $Q_{f,evap}$ and $Q_{ht,glob}$ indicate the fuel evaporation heat from SOI to SOC (J) and the heat globally exchanged by the charge with the walls over the combustion cycle (J), and $m_{f,inj}$ is the total injected fuel mass per cycle/cylinder. The experimental values of the $Q_{f,evap}$ and $Q_{ht,glob}$ parameters were identified on the basis of the measured in-cylinder pressure traces for all the tests shown in Table 2 (see [40]), and the following correlations were identified for the two parameters as a function of the engine speed, load and intake manifold temperature:

$$Q_{f,evap}[J] = 2.22E\text{-}11\, T_{int}^{3.76} n^{0.584} q_{f,inj}^{0.270} \qquad (13)$$

$$Q_{ht,glob}[J] = 117\, n^{-0.642} q_{f,inj}^{1.50} \qquad (14)$$

Estimation of the in-cylinder pressure

The in-chamber pressure was evaluated during the combustion interval using a single-zone model [41]:

$$dp = \left(\frac{\gamma - 1}{V}\right)\left(dQ_{net} - \frac{\gamma}{\gamma - 1} pdV\right) \qquad (15)$$

where the isentropic coefficient $\gamma = c_p/c_v$ was set constant and equal to 1.37. Although the physical value of γ depends on the engine operating condition, the choice of adopting a constant value of γ was done in order not to generate misalignments with the combustion metrics derived by the acquisition software of the test bench (which implements γ=1.37). An explicit integration method was used to solve Eq. (15), as it was verified that it leads to a good accuracy and is stable for crank angle integration steps ranging from 0.1 to 2 deg. Therefore, the adoption of implicit integration methods is not necessary.

Polytropic evolutions were assumed to calculate the in-cylinder pressure during the compression and expansion phases:

$$pV^m = const \qquad (16)$$

$$pV^{m'} = const \qquad (17)$$

The starting condition for the evaluation of the compression phase, i.e. the in-chamber pressure at IVC (Intake Valve Closure), was correlated to the pressure in the intake manifold p_{int}, using a correction factor Δp_{int}, as follows:

$$p_{IVC} = p_{int} + \Delta p_{int} \qquad (18)$$

The experimental values of the m, m' *and* Δp_{int} parameters were identified on the basis of the measured in-cylinder pressure traces for all the tests shown in Table 2 (reference can be made to [40] for the complete procedure). The following correlations were identified for the three parameters, which are functions of the intake manifold thermodynamic conditions and of the engine load and speed:

$$m = 3.02\, T_{int}^{-0.173} n^{0.0167} q_{f,inj}^{0.0148} \qquad (19)$$

$$m' = 2.78\, T_{int}^{-0.0919} n^{-0.0176} q_{f,inj}^{-0.0209} \qquad (20)$$

$$\Delta p_{int}[bar] = 0.163\, p_{int}^{1.06} n^{-0.000825} q_{f,inj}^{-0.0429} \qquad (21)$$

The in-cylinder pressure values during the intake and exhaust phases were considered as constants and were set equal to the pressure in the intake and exhaust manifolds, respectively [40]. It has in fact been verified that this assumption leads to a very small error in the prediction of IMEP.

The simulation of the in-cylinder pressure traces allows several combustion parameters, such as PFP and IMEP, to be evaluated.

Estimation of FMEP

The Chen-Flynn approach [43] was adopted to estimate FMEP.

The experimental values of FMEP were evaluated as the difference between the experimental values of the net IMEP and the measured values of BMEP, as follows:

$$FMEP = IMEP - BMEP \qquad (22)$$

As previously stated, the values of IMEP of one of the four cylinders were taken as being representative of all the cylinders, due to the low cylinder-to-cylinder dispersion.

The following correlation was identified to evaluate FMEP for the engine considered in this study:

$$FMEP[bar] = -1.38E\text{-}04\, n + 6.67E\text{-}08\, n^2 + \\ +0.0126\, PFP + 1.04 \qquad (23)$$

where the engine speed is expressed in rpm and PFP is expressed in bars.

The squared correlation coefficient R^2 between the predicted and experimental values of FMEP is the order of 0.89.

Model inversion

The physical model was also inverted in order to predict the value of SOI_{main} and injected fuel quantity $q_{f,inj}$ that allow given targets of MFB50 and BMEP to be achieved. Model inversion is in fact mandatory in order to use the model in a combustion control algorithm.

The procedure for model inversion is shown in [16]. The model has been inverted adopting an iterative procedure, in which the first run is based on the initial assumption of the control variables (SOI_{main} / $q_{f,inj}$) and a cycle based integral control is applied to adjust the value of control variables in order to reach the convergence of the target variables (MFB50/BMEP). The iterative procedure stops when the difference between the predicted value and required value of the target variable falls below the predefined threshold. In general, it was shown in [16] that the convergence can be achieved within two or three iterations given a reasonable level of threshold (e.g., 0.2 degree for MFB50 and 0.1 bar for BMEP). The controller is able to increase or decrease its own integral gain according to the instantaneous error so as to accelerate the convergence.

NNRT model

Direct NNRT model to estimate MFB50 and BMEP

A neural network-based real-time (NNRT) model has been proposed in this paper to estimate MFB50 and BMEP. This model is constituted by ANNs, which have been trained using data simulated by the physical engine model. The NNRT model is characterized by the same inputs as the physical engine model, namely engine speed, rail pressure, pressure and temperature in the intake manifold, O_2 concentration in the intake manifold, pressure in the exhaust manifold, injection timing and quantity of each pulse. More in detail, the proposed NNRT model has been realized through single hidden layer feed-forward neural networks. Moreover, it was verified that the use of 10 neurons is sufficient to guarantee a good prediction accuracy of the model (see the next sub-sections). Thanks to the relatively simple mathematical structure, the computational time required by this approach is much smaller than that of the physical model, and therefore it is more suitable for the development of model-based combustion control algorithms to be implemented in the ECU.

The physical engine model has been used as a virtual engine to perform virtual experiments, in order to produce the training dataset for the NNRT model. The test plan was obtained through a DoE (Design of Experiment) methodology, in which the main engine operating parameters were varied. A 3.4 GHz PC equipped with 32 GB of RAM was used to perform the simulations.

Two approaches for the generation of DoE tests have been investigated in order to train the NNRT model.

1. Global DoE. The test plan was generated using the latin hypercube sampling method. The variation range of each variable was derived from the maximum and minimum values identified over the entire experimental engine map tests (see Fig. 1). This approach has been referred to as “global DoE” because engine load (i.e., injected fuel quantity) and speed have been included among the input variables of the DoE. Therefore, a single DoE has been generated. The total number of tests of the global DoE was 22000. To generate these tests, the simulations carried out with the physical model required about 200 seconds. The NNRT model generated with this approach will be referred as GNN10 (10 indicates the neuron number).

2. Local DoE. First, a 2D map was generated for each input variable of the DoE as a function of the engine speed and injected fuel quantity, on the basis of the experimental engine map tests (see Fig. 1). These maps basically are used to store the nominal values of the main engine variables (e.g., rail pressure, boost pressure, etc.) as a function of engine load and speed. Then, a subset of 111 engine operating conditions (in terms of engine load and speed) has been selected. These operating conditions are referred to as “key-points”, and are shown in Fig. 2 in terms of injected fuel quantity and engine speed. A local DoE has then been generated for each key-point, using the Latin hypercube sampling method, by varying the input variables around the reference values (given by the previously developed 2D maps) in a range of ±20%. The complete training dataset is the assembly of all the local DoE test plans. This method has been referred to as “local DoE” because the training dataset is given by the combination of several local DoEs, each one generated at a given key-point. Therefore, engine speed and load have not been used as input parameters of the DoE, as in case 1. The total number of tests of the local DoE was chosen to be of the same order of that of the global DoE, and resulted to be equal to 22200 (i.e., 200 tests for each of the 111 key-points). To generate these tests, the simulations carried out with the physical model required a similar amount of time to that required by the global DoE (i.e., about 200 s). The NNRT model generated with this approach will be referred as LNN10.

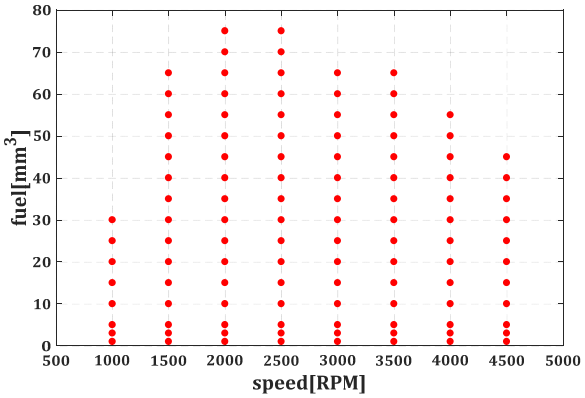


Figure 2. Key-points of the local DoE

The input variables of the DoEs are listed in Table 3.

Table 3. Input variables of the DOE

p_f	Rail pressure
p_{int}	Intake manifold pressure
T_{int}	Intake manifold temperature

O_2	Intake oxygen concentration
pexh	Exhaust manifold pressure
DT_{pil1}, DT_{pil2}	Dwell time of first and second pilot injections
SOI_{main}	SOI of main injection
qpil1, qpil2	Fuel quantity of first and second pilot injections
qf,inj	Total fuel quantity (only for the global DoE)
n	Engine speed (only for the global DoE)

Inverted NNRT model to estimate SOI_{main} and $q_{f,inj}$ from given targets of MFB50 and BMEP

In addition to the NNRT model that estimates BMEP and MFB50, an inverted NNRT model has also been developed. Similarly to the inverted physical model, the inverted NNRT model is able to predict the values of SOI_{main} and injected fuel quantity $q_{f,inj}$ in order to achieve specific targets of MFB50 and BMEP. The inverted physical model was used to generate the training dataset for the inverted NNRT model.

The inverted NNRT model is suitable to realize a model-based control of MFB50 and BMEP. It should be noted that the inverted NNRT model does not require any iteration to be carried out.

Influence of neuron number

A sensitivity analysis has been made to investigate the influence of the neuron number on the prediction accuracy of the NNRT model. The local DoE has been chosen as the standard training approach, on the basis of the results shown in the previous section, and the tested neuron numbers are equal to 5, 10, 15 and 20.

The trained NNRT model has been tested in terms of MFB50/BMEP prediction (direct approach) and $SOI_{main}/q_{f,inj}$ prediction (inverted approach), for the steady state tests reported in Tab. 2. The results are reported in Table 4, in which one can find the RMSE (Root Mean Squared Error) values, with reference to the DoE tests (Tab. 4a) and the engine map tests (Tab. 4b) shown in Tab. 2.

Table 4. Neuron number influence on the accuracy of the NNRT model, for the DoE (a) and engine map tests (b). The RMSE values of MFB50, BMEP, SOI_{main} and $q_{f,inj}$ are reported.

Model type (direct/inverted)	Predicted variable	LNN5	LNN10	LNN15	LNN20
Direct	MFB50	1.63 deg	1.18 deg	1.13 deg	1.15 deg
Direct	BMEP	0.41 bar	0.24 bar	0.16 bar	0.19 bar

(a) DoE tests: RMSE values

Model type (direct/inverted)	Predicted variable	LNN5	LNN10	LNN15	LNN20
Direct	MFB50	1.12 deg	1.10 deg	1.06 deg	1.05 deg
Direct	BMEP	0.65 bar	0.51 bar	0.52 bar	0.44 bar
Inverted	SOI_{main} (MFB50 inversion)	1.39 deg	1.19 deg	1.11 deg	1.04 deg
Inverted	$q_{f,inj}$ (BMEP inversion)	1.49 mm^3	1.45 mm^3	1.43 mm^3	1.43 mm^3

(b) engine map tests: RMSE values

With reference to the prediction of MFB50 and BMEP, a large difference can be observed when the NNRT model has been trained with 5 neurons or 10 neurons respectively, especially for the DoE tests. When the neuron number is higher than 10, the prediction accuracy does not vary remarkably.

The same behavior can be seen for the inverted NNRT model, with specific reference to the inversion of MFB50 (Tab. 4b, 3rd row). In this case, the RMSE values are similar for LNN10, LNN15 and LNN20 models (they are around 1.1 deg), while the accuracy of the LNN5 model is worse (RMSE=1.4 deg).

With reference to the inversion of BMEP (Tab. 4b, 4th row) The accuracy in the estimation of $q_{f,inj}$ is not significantly influenced by the neuron number, as the RMSE values range between 1.4 and 1.5 mm^3.

These results indicate that a 10-neuron NNRT model is sufficient to guarantee a good prediction accuracy. NNRT models characterized by a lower number of neurons suffer from underfitting, and the resulting prediction accuracy is poor. The adoption of a higher number of neurons than 10 does not lead to any improvement in the prediction accuracy. On the contrary, the adoption of a high number of neuron may lead to an excessively high number of degrees of freedom, with a consequent risk of overfitting. As a consequence, 10 neuron numbers have been used and the performance of LNN10 and GNN10 models will be evaluated in the next sections.

Results and discussion

The main results that will be reported in this section can be summarized as follows. First, the performance of the newly proposed GNN10 and LNN10 models and of the physical model used for training will be evaluated at both steady-state conditions and transient conditions, over a WLTP cycle. Both direct and inverted model versions will be considered. It will be shown that the LNN10 model provides a better accuracy than the GNN10 one and that its accuracy is similar to that of the physical model. Then, it will be shown that the approach proposed in this paper (i.e., the training of the NNRT model on the basis of a dataset generated by the physical model, used as a virtual engine) is better than a direct training of the neural networks using the bench data (the latter approach will be referred to as 'BNN'). In particular, it will be demonstrated that even when using all available experimental data (about 1000) for BNN training, the performance of the BNN model is lower than that of the LNN10 one over WLTP cycle. Viceversa, the performance of the LNN10 model is still good when the physical model used for training is calibrated only using 25 experimental data. It is also worthwhile recalling that the virtual dataset for the LNN10 training includes 22200 simulated points and requires only 200 s to be generated. Therefore, the LNN10 approach requires very low experimental effort and low computational effort to be trained. Finally, a comparison between the computational time required by the physical model and that required by the LNN model will be reported.

Model application: steady state tests

First, the physical and NNRT models have been applied to the steady state tests reported in Tab. 2. The investigated models are the physical engine model (referred to as 'physical' in the subsequent figures), the GNN10 model and the LNN10 model. Figures 3 and 4 report the predicted vs. experimental values of MFB50 and BMEP, respectively, using the three abovementioned models, for the DoE tests and engine map tests reported in Tab. 2. The prediction accuracy of each model has been quantified by the squared correlation coefficient (R^2) and by the RMSE, which are reported in each figure.

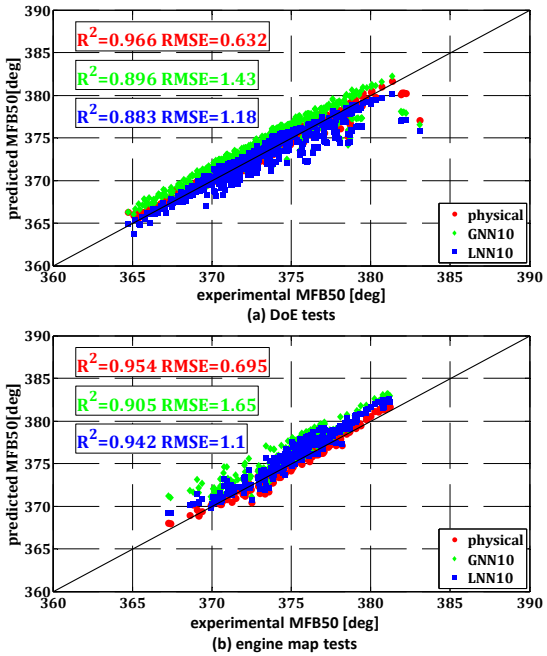


Figure 3. Predicted vs. experimental MFB50 values for the steady-state tests reported in Tab. 2.

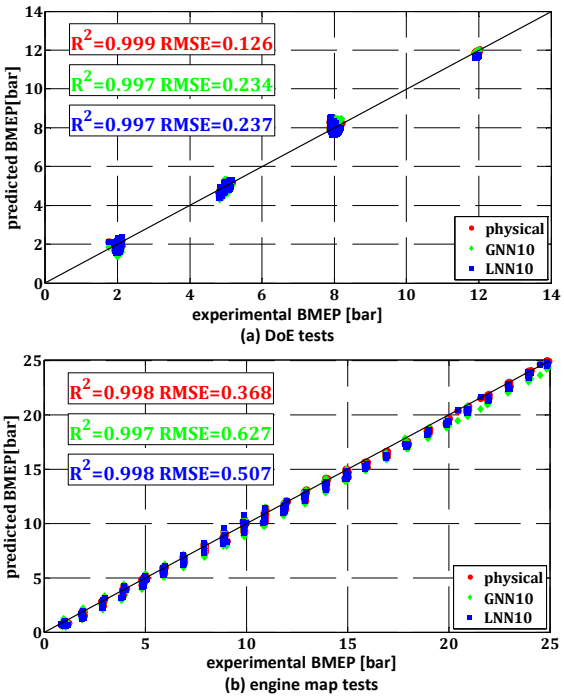


Figure 4. Predicted vs. experimental BMEP values for the steady-state tests reported in Tab. 2.

With reference to MFB50 prediction, it can be seen that the RMSE of the physical model is between 0.6 and 0.7 for both the DoE and engine map tests. The accuracy of the LNN10 model is lower than that of the physical model. However, the RMSE values are still satisfactory, as they are of the order of 1.1 deg for both test categories. The performance of the GNN10 model is worse than that of the LNN10 one, since it is characterized by RMSE values which are around 1.5 deg.

With reference to the prediction of BMEP, the RMSE of the physical model is around 0.11 bar and 0.37 bar for the DoE tests and engine map tests, respectively. The LNN10 model is characterized by higher RMSE values than the physical model, which are however satisfactory (0.237 bar and 0.507 bar for the two test categories, respectively). The performance of the GNN10 model is similar to that of the LNN10 model for the DoE tests, but it is worse for the engine map tests (RMSE=0.627 bar).

Based on the prediction accuracy of these three different approaches, one can conclude that the physical model is still the most accurate one because its physical consistency assures that MFB50 and BMEP can be well captured in most cases. The NNRT model is less accurate than the physical model, however the accuracy is still good for control-oriented applications, especially if the local DoE strategy is adopted for the training. The adoption of the local DoE strategy, in particular, allows an improvement in the MFB50 prediction to be obtained, compared to the global DoE strategy. The main reason of this behavior is that, for the global DoE approach, there is no physical constraint in the process of virtual test generation. Thus, some tests can be generated which actually are not feasible in the real engine (e.g., high load operating conditions with low values of injection pressure or low values of intake oxygen concentration). The inclusion of this kind of tests in the training dataset of the NNRT model can affect the resulting model accuracy. On the other hand, for the local DoE approach, since the virtual tests are generated by means of local

DoEs around the nominal working conditions, the physical compatibility between input parameters is guaranteed, and most of the generated tests can be considered as physically consistent. Therefore, the local DoE approach is more effective in generating valid virtual tests compared with the global DoE method, and consequently the resulting training procedure of NNRT model is more reliable.

The models have also been inverted to estimate the values of SOI_{main} and $q_{f,inj}$ in order to achieve targets of MFB50 and BMEP. The results have been reported in Figure 5 and Figure 6 for MFB50 and BMEP, respectively. The accuracy indices (R^2 and RMSE) have also been reported in each figure.

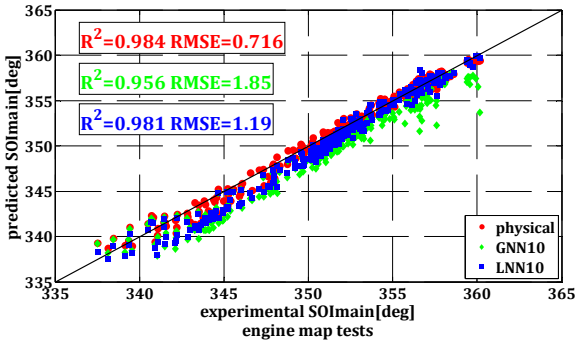


Figure 5. Inverted models: predicted vs. experimental values of SOI_{main} for the engine map tests reported in Tab. 2

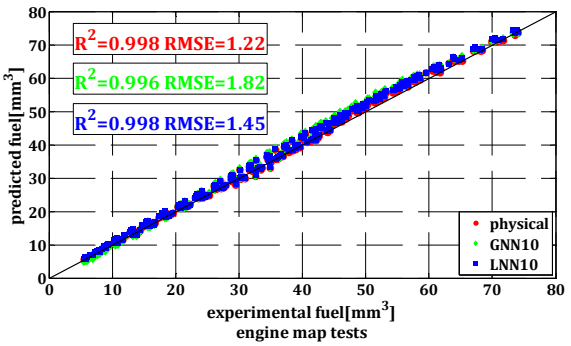


Figure 6. Inverted models: predicted vs. experimental values of $q_{f,inj}$ for the engine map tests reported in Tab. 2

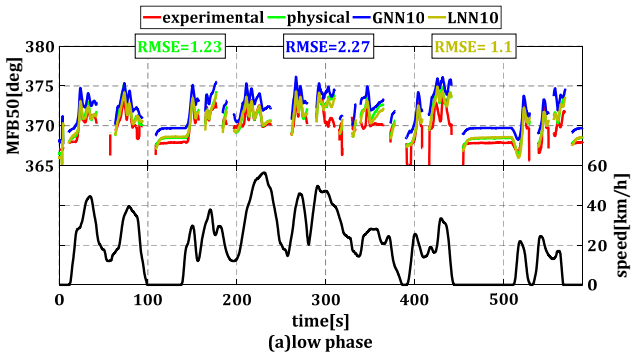
With reference to MFB50 inversion, the physical model has the lowest RMSE (0.7 deg). The RMSE of the LNN10 model is slightly higher (1.19 deg), while the GNN10 model shows the worst accuracy (RMSE=1.8 deg).

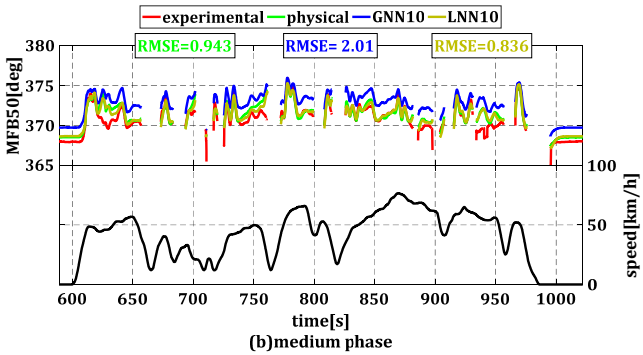
With reference to the inversion of BMEP, the accuracy of the physical model and NNRT model based on local DoE strategy (LNN10) is quite similar, with RMSE values around 1.2 mm^3 and 1.4 mm^3, respectively. The performance of the GNN10 model is worse, since the RMSE value is around 1.8 mm^3.

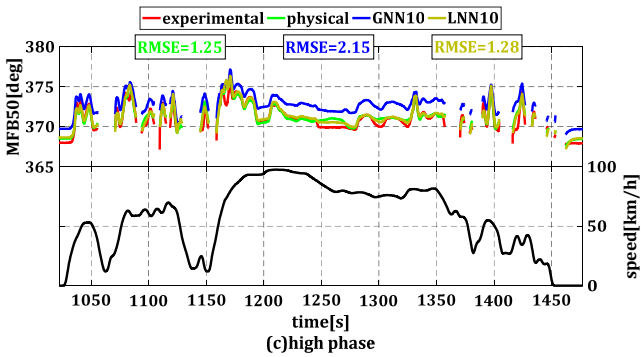
It should also be noted that a combined strategy, in which the NNRT model was trained using both the global DoE and local DoE methods at the same time, has been tested. It was found that the accuracy of the model is better than that of the LNN10 model in terms of MFB50 prediction, especially for the DoE tests, but worse in terms of BMEP prediction.

Model application: WLTP cycle

The physical and NNRT models have then been applied in transient conditions over the WLTP cycle. Figures 7 and 8 report the predicted and experimental trends of MFB50 and BMEP, respectively. Each curve is identified by color, as indicated in the legend. The RMSE values of each model have been reported in the charts, and the text color is coherent with that of the legend. The reported RMSE values are related to the physical model (green, on the left), to the GNN10 NNRT model (blue, in the middle) and to the LNN10 NNRT model (yellow, on the right), respectively. The WLTP cycle has been split into 4 phases, which have been shown in separate charts. The vehicle speed profile is also shown at the bottom of each figure. It should be noted that the GNN10 model has also been considered in this analysis, in order to verify its performance in transient conditions. It should be noted that, in the charts, MFB50 is reported only for positive values of BMEP, since it is not meaningful when BMEP is negative.







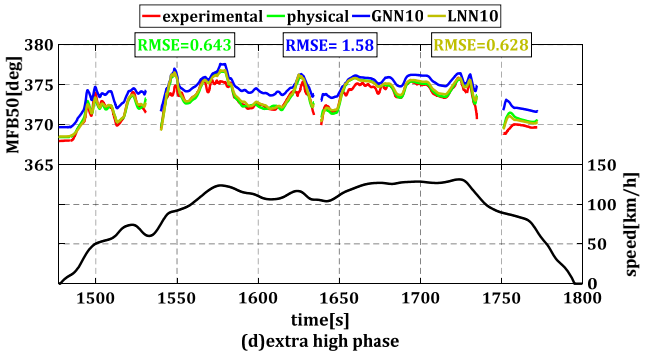


Figure 7. Experimental and predicted trends of MFB50 over the WLTP cycle using the physical and NNRT models.

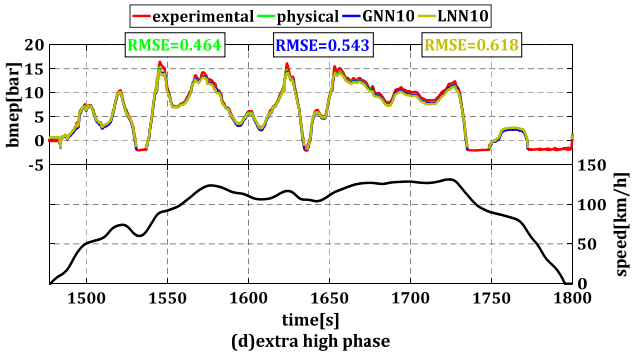
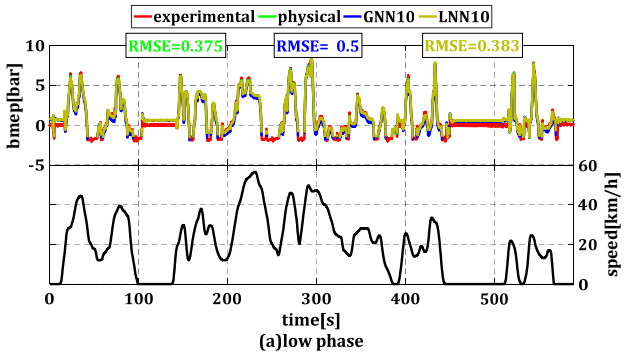
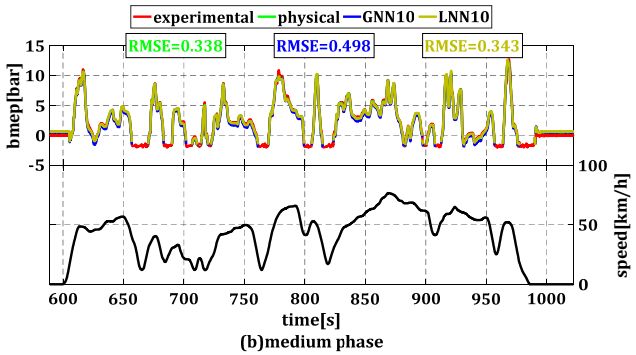
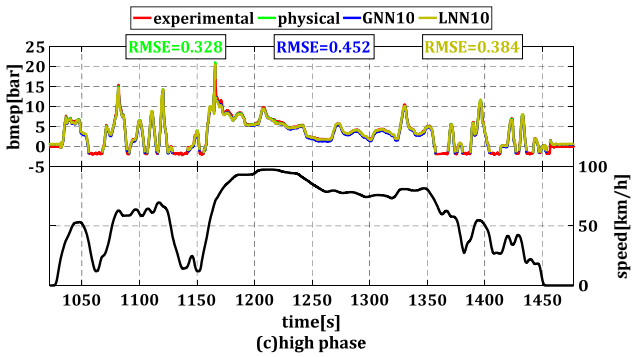


Figure 8. Experimental and predicted trends of BMEP over the WLTP cycle using the physical and NNRT models.





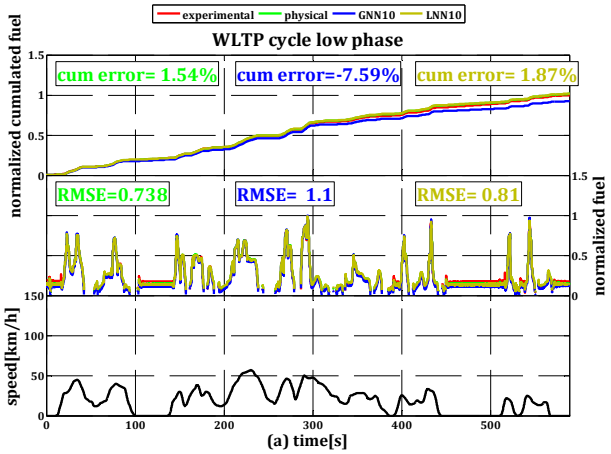


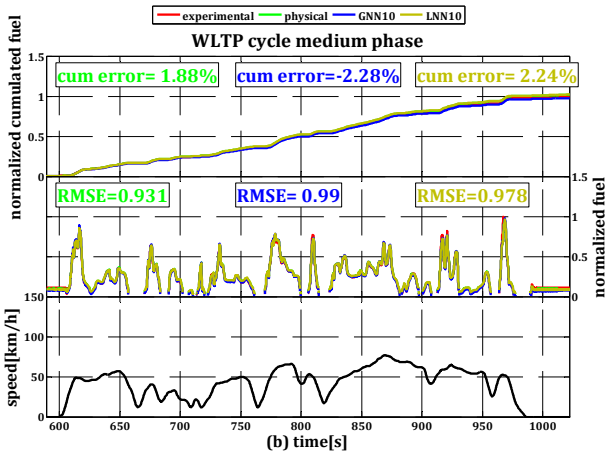
With reference to MFB50 prediction, it can be seen in Fig. 7 that the LNN10 NNRT model has a similar performance to that of the physical model, with RMSE values around 1.1 deg in the low, medium and high phases, and RMSE values around 0.6 deg in the extra-high phase. The RMSE values of the GNN10 model is much higher compared with that of the physical and LNN10 models.

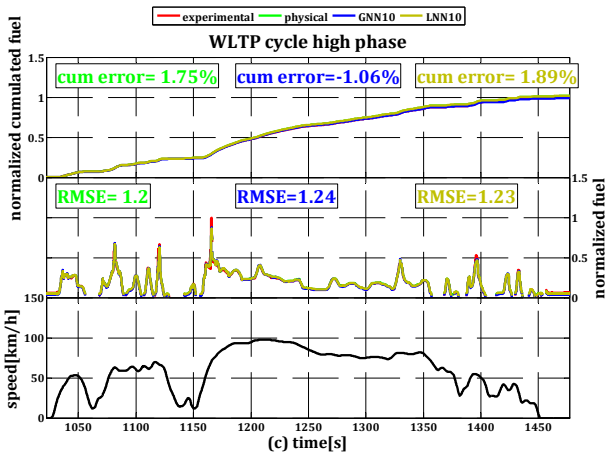
With reference to BMEP prediction (Fig. 8), a similar behavior can be observed. In this case, the physical model and the LNN10 NNRT model predict BMEP with a similar accuracy (RMSE values range between 0.3 and 0.4 bar for the first three cycle phases). In the extra-high phase, the RMSE value of LNN10 is slightly higher than that of the physical model (0.6 bar vs. 0.45 bar).

The results of the application of the NNRT models over the WLTP cycle confirm that the adoption of the local DoE training strategy leads to a higher predictive capability than the adoption of the global DoE strategy. This conclusion is in line with the results that were shown for the steady state tests.

The results of the model inversion have also been reported, for BMEP only, in Fig. 9, which displays the experimental and predicted trends of cumulative fuel consumption (top) and instantaneous fuel consumption (middle). The values have been normalized for confidentiality reasons. The RMSE values of instantaneous fuel consumption and the percentage error of cumulative fuel consumption at the end of each phase are also reported in the charts. The color of the lines and of the performance indices are specified by the legend.







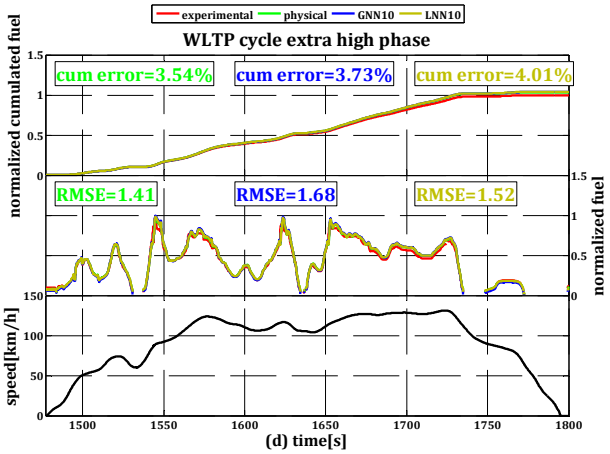


Figure 9. Experimental and predicted trends of fuel consumption over the WLTP cycle using the inverted physical and NNRT models.

These results show that all the three models are able to estimate the fuel consumption properly with percentage error at the phase end lower than 5%. However, in the low phase of WLTP the error of the GNN10 is significantly higher than the error of the other two models, as the fuel consumption at the end of that phase has been underestimated by 7%. Therefore, the LNN 10 results to be the best method also for BMEP model inversion in transient conditions.

The performance of the inverted NNRT models in terms of MFB50 could not be verified over the WLTP, since the cycle-by-cycle experimental values of MFB50 (to be used as targets) were not available.

Training of NNRT model: experimental dataset vs. virtual dataset

In the previous sections, it has been shown that the NNRT model trained using a virtual dataset generated by the physical engine model is capable of providing a good level of accuracy in the prediction of BMEP and MFB50, at both steady-state and transient conditions. The main advantage of this approach is represented by the fact that the generation of virtual tests for the NNRT model training does not require any experimental cost. In addition, a small amount of time is needed for the generation of the virtual plan (about 200 s to generate 22200 data), since the physical model used for the generation of the virtual dataset is of the low-throughput type (see [16]). At the same time, the physical engine model used for the NNRT model training requires a limited number of experimental tests to be calibrated thanks to its physical consistency (see [16]), therefore the required overall experimental effort (and related cost) to train the NNRT models is low. In short, the NNRT model trained with the virtual dataset is able to reproduce the results of the physical engine model with a good level of accuracy. However, it is characterized by a much lower computational time, so that it is more suitable to be implemented in the ECU for real-time control tasks.

On the contrary, a neural network-based model calibrated using experimental tests would require a large number of test conditions to be properly trained, in order to guarantee a good predictive capability at both steady-state and transient conditions.

In this section, a comparison will be made between the performance of a neural network-based model trained using the 985 available bench tests shown in Table 2 (this model will be referred to as 'BNN') and the performance of the LNN10 model trained using the virtual dataset and the local DoE approach (the training is based on 22200 virtual tests, as explained in the previous sections, using a physical model calibrated on the experimental 985 tests). The training strategy based on experimental tests will be referred to as "bench test-based training strategy" in what follows.

Different neuron numbers have been considered for the BNN model, namely 1, 2, 3, 5 and 10. It should be recalled that a low number of neurons may result in model underfitting, while a high number of neurons may instead result in model overfitting. In general, the number of free coefficients of the considered neural network is equal to:

$$N_{coeff} = N_n \cdot 15 + 2 \qquad (24)$$

where N_n is the neuron number. The performance of the BNN model has been tested over the steady state tests reported in Table 2 and the results are reported in Table 5 and Table 6 for the DoE tests and engine map tests, respectively. It can be seen that, for the BNN model, the number of neurons has been varied in a range between 1 and 10 (the corresponding models are referred to as BNN1, …, BNN10). The results of the LNN10 model have also been reported for comparison.

Table 5. Prediction accuracy of BNN and LNN10 models for the DoE tests (a) and engine map test (b). The RMSE values of MFB50 and BMEP, obtained using the different models, are reported in the table.

	BNN1	BNN2	BNN3	BNN5	BNN10	LNN10
MFB50	3.23 deg	1.13 deg	0.49 deg	0.29 deg	0.11 deg	1.13 deg
BMEP	0.76 bar	0.17 bar	0.18 bar	0.13 bar	0.05 bar	0.16 bar

(a) DoE tests: RMSE values

	BNN1	BNN2	BNN3	BNN5	BNN10	LNN10
MFB50	2.62 deg	1.37 deg	0.57 deg	0.29 deg	0.15 deg	1.06 deg
BMEP	0.67 bar	0.28 bar	0.27 bar	0.20 bar	0.09 bar	0.52 bar

(b) engine map tests: RMSE values

The results shown in Table 5 refer to the case in which the BNN models have been trained and applied to the same experimental dataset, i.e., that shown in Table 2. It can be observed in the table that the BNN models are characterized by very small values of RMSE for both MFB50 and BMEP prediction, and the accuracy increases with the number of neurons. However, this does not necessarily mean that the BNN models are robust, since there is a risk of overfitting when the number of neurons increases. In order to verify this, the BNN models have then been applied to the WLTP cycle, and the results are shown in Table 6, which reports the values of RMSE of MFB50 and BMEP for each phase of the WLTP cycle and over the entire cycle.

Table 6. Prediction accuracy of the BNN and LNN10 models over the WLTP cycle. The RMSE values of MFB50 and BMEP are reported for the different phases of the WLTP cycle.

	low phase	medium phase	high phase	extra high phase	overall
BNN1	2.92 deg	2.22 deg	2.48 deg	1.36 deg	2.39 deg
BNN2	1.51 deg	1.28 deg	1.45 deg	1.48 deg	1.44 deg
BNN3	4.00 deg	2.23 deg	1.91 deg	1.76 deg	2.77 deg
BNN5	5.01 deg	3.79 deg	2.37 deg	1.27 deg	3.59 deg
BNN10	6.54 deg	4.22 deg	3.19 deg	2.40 deg	4.60 deg
LNN10	1.11 deg	0.84 deg	1.28 deg	0.63 deg	1.02 deg

(a) RMSE values of MFB50

	low phase	medium phase	high phase	extra high phase	overall
BNN1	0.36 bar	0.49 bar	0.61 bar	0.48 bar	0.49 bar
BNN2	0.71 bar	0.56 bar	0.60 bar	0.51 bar	0.61 bar
BNN3	0.69 bar	0.54 bar	0.58 bar	0.54 bar	0.60 bar
BNN5	0.85 bar	0.60 bar	0.58 bar	0.57 bar	0.68 bar
BNN10	0.92 bar	0.83 bar	0.77 bar	0.59 bar	0.81 bar
LNN10	0.38 bar	0.34 bar	0.38 bar	0.62 bar	0.43 bar

(b) RMSE values of BMEP

It can be observed that the RMSE values of the BNN models are evidently higher than those of LNN10, regardless of neuron number. With reference to MFB50, the best accuracy is obtained with a number of neurons equal to 2, even though the resulting RMSE values are higher than those of the LNN10. These results clearly indicate that the BNN models are not sufficiently robust when applied over transient conditions, even though they are calibrated over about 1000 experimental tests. An improvement in the performance of the BNN would therefore require an even larger number of experimental tests, in order to cover the actual engine operating conditions which occur over the transient conditions (and which may be largely different from the steady-state conditions, due to effects such as turbo-lag, flow delays in the pipes, etc.).

On the contrary, the LNN10 model, which is based on a large virtual training dataset generated by a physical model, is much more robust when applied in transient operation, and it is capable of providing a similar level of accuracy to that of the physical model.

Model training with limited number of experimental tests

It is well known that carrying out the experimental activity at the test bench is time consuming and expensive. In the previous section, it has been shown that the training of a neural network-based model based on experimental tests (i.e., BNN model) would require a very large number of tests in order to be sufficiently robust also in transient conditions. On the contrary, the NNRT model proposed in this study is trained on the basis of a virtual dataset, produced by a physical low-throughput engine model. It has already been shown in

[16] that the physical engine model features high robustness and can be properly calibrated with very small number of experimental tests. Therefore, in the end, a robust NNRT model can be obtained even using a small number of experimental tests.

This has been further verified in this section. In particular, the physical engine model has been recalibrated with 2.5% of experimental tests reported in table 2 (i.e., 25 tests), and the LNN10 model has been retrained on the basis of a virtual dataset generated by the physical model recalibrated with 25 tests. The prediction accuracy of the derived NNRT model in steady state tests is reported in Table 7 (column indicated with 'LNN10 (2.5%)'), in which the percentage of experimental data used for the physical model calibration has been specified in brackets. The same amount of experimental tests has also been used to directly train the BNN model (column indicated with 'BNN1 (2.5%)'. In this case only 1 neuron could be used because the number of free coefficients of the neural network has to be lower than the number of tests used for training, in order to make the problem well defined. The BNN1 is characterized by 17 free coefficients (see Eq. (24)), so it is the only suitable choice when using 25 experimental tests.

Table 7. Prediction accuracy of BNN1 and LNN10 models when using only 2.5% of experimental tests reported in table 2. The accuracy of the LNN10 model obtained using 100% data for the physical model calibration is also reported for reference. The RMSE values of BMEP and MFB50 are reported.

	BNN1 (2.5%)	LNN10 (2.5%)	LNN10 (100%)
MFB50	3.55 deg	1.15 deg	1.18 deg
BMEP	0.80 bar	0.34 bar	0.24 bar

(0) DoE tests: RMSE values

	BNN1 (2.5%)	LNN10 (2.5%)	LNN10 (100%)
MFB50	2.64 deg	1.21 deg	1.10 deg
BMEP	0.81 bar	0.53 bar	0.51 bar

(b) engine map tests: RMSE values

It can be seen that the LNN10 model trained using a virtual dataset produced by a physical model calibrated with 2.5% of available tests has a similar level of accuracy to the case in which the physical model has been calibrated with 100% of available tests (column indicated with 'LNN10 (100%)'). On the other hand, much larger errors are observed for the BNN1 model calibrated with 2.5% of available experimental data.

Finally, it should be noted that for the LNN10 model there is virtually no limit in the dimension of the training data set, because these data are obtained from the simulations carried out with the physical model. Therefore, the neuron number can be freely chosen in order to accurately reproduce the performance of the physical model. On the contrary, the choice of the neuron number of the BNN model is constrained by the number of available experimental tests. This constraint becomes more critical when a small number of tests is used for training.

Computational time

The physical and NNRT models have also been compared in terms of computational time. In particular, the NNRT model has been tested with different neuron numbers to investigate the corresponding influence. The models have been run on a PC with CPU frequency 3.4 GHz and RAM 32 GB using Matlab. A crank angle computational step of 0.1 deg was used for the physical model. The result are shown in Table 8.

Table 8. Computational time (ms) required by the LNN and physical models for MFB50/BMEP prediction

LNN5	0.095 ms
LNN10	0.098 ms
LNN15	0.095 ms
LNN20	0.096 ms
physical	1.534 ms

The data reported in Table 8 clearly show that the NNRT model is much more computationally efficient, as the required computational time is less than 10% of the physical model. Moreover, the neuron number does not influence the computational time required by the NNRT model.

The advantage of the NNRT model in this aspect is even larger when the model inversion is required. It has been shown in [16] that, for physical model in general two or three iterations are necessary to achieve the convergence and estimate the values of SOI_{main} and $q_{f,inj}$ which are necessary to achieve specific targets of MFB50 and BMEP. Instead, the inverted NNRT model does not require any iteration, as the model is trained using the inverted physical model, so that MFB50 and BMEP targets are given as input, while SOI_{main} and $q_{f,in}$ are model outputs. Therefore, in the case of model inversion, the computational time required by the inverted NNRT model can be less than 5% of that of the inverted physical model.

Summary/Conclusions

A real-time approach has been developed and assessed to control BMEP (brake mean effective pressure) and MFB50 (crank angle at which 50% of fuel mass has burnt) in a Euro 6 1.6L GM diesel engine. The approach is based on the use of feed-forward ANNs (artificial neural networks), which have been trained using virtual tests simulated by a previously developed low-throughput physical engine model. The resulting model, which is referred to as NNRT (Neural Network-based Real Time), therefore, has the aim of replicating the outcomes of the physical engine model, but with a much shorter computational time, so that it can be implemented onboard for model-based combustion control tasks. In particular, the time required by NNRT model is less than 10% of physical model, with a similar level of accuracy

Two different training approaches have been investigated. The first one is based on a global DoE in which all the main engine variables (including load and speed) are varied (this approach was referred to as 'GNN10', since it is based on 10 neurons), while the second one is

based on local DoEs generated at specific engine key-points (this approach was referred to as 'LNN10'). A sensitivity analysis has been implemented to investigate the influence of neuron number on the model accuracy. It has been shown that 10 neurons are sufficient to obtain a satisfactory accuracy of the NNRT model, which is in line with that of the physical model also in transient operation. A higher number of neurons in fact does not improve the model performance

Both the physical and NNRT models have been inverted in order to estimate the required SOI_{main} (start of injection of the main pulse) and injected fuel quantity $q_{f,inj}$ to achieve specific targets of MFB50 and BMEP. Model inversion is in fact mandatory for the development of a combustion control algorithm.

The performance of the models has been evaluated at steady state and transient conditions, over the WLTP cycle. It has been found that, in general, the LNN10 model shows better performance than the GNN10 one, especially for the prediction of MFB50.

Moreover, it has been verified that a neural-network approach trained using experimental data (this approach was referred to as 'BNN') is less robust than the proposed LNN10 model when applied over WLTP cycle, even when it is calibrated using all available experimental data (about 1000). On the contrary, the proposed LNN10 model is still robust when the physical model used for training is calibrated only using 25 experimental data. In general, the virtual dataset for the LNN10 training includes 22200 points and requires only 200 s to be generated, since the physical model is of a low-throughput type. In conclusion, the LNN10 approach requires very low experimental effort and low computational effort to be trained, and is robust when applied at both steady-state and transient conditions.

References

1. Fontanesi, F., Giacopini, M., "Multiphase CFD-CHT optimization of the cooling jacket and FEM analysis of the engine head of a V6 diesel engine", *Applied Thermal Engineering* 52 : 293-303, 2013, doi: 10.1016/j.applthermaleng.2012.12.005.
2. Jemni, M.A., Kantchev, G., Abid, M.S., "Influence of intake manifold design on in-cylinder flow and engine performances in a bus diesel engine converted to LPG gas fuelled, using CFD analyses and experimental investigations.", *Energy* 36:2701-2715, 2011, doi: 10.1016/j.energy.2011.02.011.
3. Lee, C.H., Reitz, R.D., "CFD simulations of diesel spray tip penetration with multiple injections and with engine compression ratios up to 100:1", *Fuel* 111:289–297, 2013, doi: 10.1016/j.fuel.2013.04.058.
4. Perini, F., Dempsey, A., Reitz, R., Sahoo, D. et al., "A Computational Investigation of the Effects of Swirl Ratio and Injection Pressure on Mixture Preparation and Wall Heat Transfer in a Light-Duty Diesel Engine", SAE Technical Paper 2013-01-1105, 2013, doi:10.4271/2013-01-1105.
5. Rakopoulos, C.D., Kosmadakis, G.M., Pariotis, E.G., "Investigation of piston bowl geometry and speed effects in a motored HSDI diesel engine using a CFD against a quasi-dimensional model", *Energy Conversion and Management* 51:470–484, 2010, doi: 10.1016/j.enconman.2009.10.010.
6. Mattarellia, E., Rinaldinia, C.A., Golovitchevb, V.I., "CFD-3D Analysis of a Light Duty Dual Fuel (Diesel/Natural Gas) Combustion Engine", *Energy Procedia* 45:929 – 937, 2014, doi: 10.1016/j.egypro.2014.01.098.
7. Mobasheri, R., Peng, Z., Mirsalim, S.M., "Analysis the effect of advanced injection strategies on engine performance and pollutant emissions in a heavy duty DI-diesel engine by CFD modeling", *International Journal of Heat and Fluid Flow* 33:59–69, 2012, doi: 10.1016/j.ijheatfluidflow.2011.10.004.
8. Jayashankara, B., Ganesan, V., "Effect of fuel injection timing and intake pressure on the performance of a DI diesel engine – A parametric study using CFD", *Energy Conversion and Management* 51:1835–1848, 2010, doi: 10.1016/j.enconman.2009.11.006.
9. Cipolla, G., Vassallo, A., Catania, A., Spessa, E. et al., "Combined application of CFD modeling and pressure-based combustion diagnostics for the development of a low compression ratio high-performance diesel engine" SAE Technical Paper 2007-24-0034, 2007, doi:10.4271/2007-24-0034.
10. Lešnik, L., Iljaz, J., Hribernik, A., Kegl, B., "Numerical and experimental study of combustion, performance and emission characteristics of a heavy-duty DI diesel engine running on diesel, biodiesel and their blends", *Energy Conversion and Management* 81:534–546, 2014, doi: 10.1016/j.enconman.2014.02.039.
11. Kéromnès, A., Delaporte, B., Schmitz, G., Le Moyne, L., "Development and validation of a 5 stroke engine for range extenders application", *Energy Conversion and Management* 82:259–267, 2014, doi: 10.1016/j.enconman.2014.03.025.
12. Arcidiacono, M., Baratta, M., Finesso, R., Kheshtinejad, H. et al., "Use of an Innovative Predictive Heat Release Model Combined to a 1D Fluid-Dynamic Model for the Simulation of a Heavy Duty Diesel Engine," *SAE Int. J. Engines* 6(3):1566-1579, 2013, doi:10.4271/2013-24-0012.
13. Baratta, M., Finesso, R., Misul, D., and Spessa, E., "Comparison between Internal and External EGR Performance on a Heavy Duty Diesel Engine by Means of a Refined 1D Fluid-Dynamic Engine Model," SAE Int. J. Engines 8(5):1977-1992, 2015, doi:10.4271/2015-24-2389
14. Montenegro, G., Onorati, A., Piscaglia, F., D'Errico, G., "Integrated 1D-MultiD Fluid Dynamic Models for the Simulation of I.C.E. Intake and Exhaust Systems", SAE Technical Paper 2007-01-0495, 2007, doi: 10.4271/2007-01-0495.
15. Rakopoulos, C. and Giakoumis, E., "Review of Thermodynamic Diesel Engine Simulations under Transient Operating Conditions," SAE Technical Paper 2006-01-0884, 2006, doi:10.4271/2006-01-0884.
16. Finesso, R., Spessa, E., and Yang, Y., "Development and Validation of a Real-Time Model for the Simulation of the Heat Release Rate, In-Cylinder Pressure and Pollutant Emissions in Diesel Engines," SAE Int. J. Engines 9(1):322-341, 2016, doi:10.4271/2015-01-9044.
17. Gani, E. and Manzie, C., "Indicated Torque Reconstruction from Instantaneous Engine Speed in a Six-cylinder SI Engine Using Support Vector Machines," SAE Technical Paper 2005-01-0030, 2005, doi:10.4271/2005-01-0030.
18. Hiroyasu, H., Miao, H., Hiroyasu, T., Miki, M. et al., "Genetic Algorithms Optimization of Diesel Engine Emissions and Fuel Efficiency with Air Swirl, EGR,Injection Timing and Multiple Injections," SAE Technical Paper 2003-01-1853, 2003, doi:10.4271/2003-01-1853.
19. Stephan Hennings Och, Luís Mauro Moura, Viviana Cocco Mariani, et al., "Volumetric efficiency optimization of a single-cylinder D.I. diesel engine using differential evolution algorithm", applied thermal engineering 108:660-669, 2016, doi: 10.1016/j.applthermaleng.2016.07.042

20. Bolan Liu, Changlu Zhao, Fujun Zhang, et al., “Misfire detection of a turbocharged diesel engine by using artificial neural networks”, applied thermal engineering 55(1-2):26-32, 2013, doi: 10.1016/j.applthermaleng.2013.02.032.
21. Samir Saraswati, Praveen Kumar Agarwal, Satish Chand, “Neural networks and fuzzy logic-based spark advance control of SI engines”, expert systems with applications 38(6): 6916-6925, 2011, doi: 10.1016/j.eswa.2010.12.032.
22. Barakat, M. and Abdelaziz, M., "Neural Network Transmission Control," SAE Technical Paper 2016-01-0089, 2016, doi:10.4271/2016-01-0089.
23. Abdullah Uzun, “Air mass flow estimation of diesel engines using neural network”, fuel 117(A):833-838, 2014, doi: 10.1016/j.fuel.2013.09.078.
24. T. Sardarmehni, J. Keighobadi, M.B. Menhaj, H. Rahmani, “Robust predictive control of lambda in internal combustion engines using neural networks”, archives of civil and mechanical engineering 13(4):432-443, 2013, doi: 10.1016/j.acme.2013.05.003.
25. Asik, J., Peters, J., Meyer, G., and Tang, D., "Transient A/F Estimation and Control Using a Neural Network," SAE Technical Paper 970619, 1997, doi:10.4271/970619.
26. Wai Kean Yap, Tien Ho, Vishy Karri, “Exhaust emissions control and engine parameters optimization using artificial neural network virtual sensors for a hydrogen-powered vehicle”, international journal of hydrogen energy 37(10):8704-8715, 2012, doi: 10.1016/j.ijhydene.2012.02.153.
27. Abdullah Uzun, “A parametric study for specific fuel consumption of an intercooled diesel engine using a neural network”, Fuel(93) 189-199, 2012, doi: 10.1016/j.fuel.2011.11.004.
28. C. Bennett, J.F. Dunne, S. Trimby, D. Richardson, “Engine cylinder pressure reconstruction using crank kinematics and recurrently-trained neural networks”, Mechanical systems and signal processing 85:126-145, 2016, doi: 10.1016/j.ymssp.2016.07.015.
29. Vijay Manikandan Janakiraman, XuanLong Nguyen, Dennis Assanis, “Nonlinear identification of a gasoline HCCI engine using neural networks coupled with principal component analysis”, applied soft computing 13(5):2375-2389, 2013, doi: 10.1016/j.asoc.2013.01.006.
30. Yusuf Cay, “Prediction of a gasoline engine performance with artificial neural network”, Fuel 111:324-331, 2013, doi: 10.1016/j.fuel.2012.12.040.
31. Yusuf Çay, Ibrahim Korkmaz, Adem Çiçek, Fuat Kara, “Prediction of engine performance and exhaust emissions for gasoline and methanol using artificial neural network”, Energy 50:177-186, 2013, doi: 10.1016/j.energy.2012.10.052.
32. Yusuf Çay, Adem Çiçek, Fuat Kara, Selami Sağiroğlu, “Prediction of engine performance for an alternative fuel using artificial neural network”, applied thermal engineering 37:217-225, 2012, doi: 10.1016/j.applthermaleng.2011.11.019.
33. Brusca, S., Lanzafame, R., and Messina, M., "A Combustion Model for ICE by Means of Neural Network," SAE Technical Paper 2005-01-2110, 2005, doi:10.4271/2005-01-2110.
34. Brahma, I., He, Y., and Rutland, C., "Improvement of Neural Network Accuracy for Engine Simulations," SAE Technical Paper 2003-01-3227, 2003, doi:10.4271/2003-01-3227.
35. Finesso, R., Spessa, E., Yang, Y., Alfieri, V. et al., "HRR and MFB50 Estimation in a Euro 6 Diesel Engine by Means of Control-Oriented Predictive Models," SAE Int. J. Engines 8(3):1055-1068, 2015, doi:10.4271/2015-01-0879.
36. Chmela, F.G., and Orthaber, G.C., “Rate of Heat Release Prediction for Direct Injection Diesel Engines Based on Purely Mixing Controlled Combustion”, SAE Technical Paper 1999-01-0186, 1999, doi:10.4271/1999-01-0186.
37. Egnell, R., “A Simple Approach to Studying the Relation between Fuel Rate, Heat Release Rate and NO Formation in Diesel Engines”, SAE Technical Paper 1999-01-3548, 1999, doi:10.4271/1999-01-3548.
38. Ryan, T. W., Callahan, T.J., “Homogeneous Charge Compression Ignition of Diesel Fuels”. SAE Technical Paper 961160, 1996, doi:10.4271/961160.
39. Ericson, C., Westerberg, B., Andersson, M., and Egnell, R., “Modelling Diesel Engine Combustion and NOx Formation for Model Based Control and Simulation of Engine and Exhaust Aftertreatment Systems”, SAE Technical Paper 2006-01-0687, 2006, doi:10.4271/2006-01-0687.
40. Catania, A.E., Finesso, R., Spessa, E., “Predictive Zero-Dimensional Combustion Model for DI Diesel Engine Feed-Forward Control”, *Energy Conversion and Management.* 52(10):3159–3175, 2011, doi:10.1016/j.enconman.2011.05.003.
41. Heywood, J.B., “Internal Combustion Engine Fundamentals”, McGraw-Hill Intern. Editions, 1988.
42. Finesso, R., Spessa, E., “Ignition delay prediction of multiple injections in diesel engines”, Fuel 119(1):170-190, 2014, doi: 10.1016/j.fuel.2013.11.040.
43. Chen, S.K., Flynn, P.F., “Development of Single Cylinder Compression Ignition Research Engine”, SAE Technical Paper 650733, 1966. doi:10.4271/650733.

Contact Information

Prof. Ezio Spessa

IC Engines Advanced Laboratory

Dipartimento Energia, Politecnico di Torino

c.so Duca degli Abruzzi, 24 - 10129 Torino (Italy)

phone: +39-011-090.4482

ezio.spessa@polito.it.

Definitions/Abbreviations

ANN	Artificial neural network
BMEP	Brake Mean Effective Pressure
BNN	Neural network trained with bench data
CA	crank angle
CFD	Computer Fluid-Dynamics
c_p	specific heat at constant pressure
c_v	specific heat at constant volume

DEA	Differential evolution algorithm
DoE	Design of experiment
ECU	Engine Control Unit
EGR	Exhaust Gas Recirculation
EOI	end of injection
FMEP	Friction Mean Effective Pressure
GM	General Motors
GM-GPS	General Motors – Global Propulsion Systems
GNN	Neural network trained with a virtual dataset generated with global DoE approach
$\mathbf{H_L}$	lower heating value of the fuel
IMEP	Indicated Mean Effective Pressure
IVC	Intake Valve Closing
K	combustion rate coefficient
LNN	Neural network trained with a virtual dataset generated with local DoE approach
m	mass; compression phase polytropic coefficient
m'	expansion phase polytropic coefficient
$\dot{m}_{f,inj}$	fuel injection rate
MFB50	crank angle at which 50% of the fuel mass fraction has burned
n	engine rotational speed
NEDC	New European Driving Cycle
NN	Neural network
NNRT	Neural-network-based real-time model
$\mathbf{O_2}$	intake charge oxygen
	concentration
p	pressure
$\mathbf{p_f}$	injection pressure
PFP	Peak firing pressure
$\mathbf{p_{int}}$	intake manifold pressure
pil	pilot injection
q	injected fuel volume quantity
$\mathbf{Q_{ch}}$	chemical heat release
$\mathbf{Q_{f,evap}}$	energy associated with fuel evaporation
$\mathbf{Q_{fuel}}$	chemical energy associated with the injected fuel
$\mathbf{Q_{ht,glob}}$	global heat transfer between the charge and the walls
$\mathbf{Q_{net}}$	net heat release
$\mathbf{q_{f,inj}}$	total injected fuel volume quantity
$\mathbf{q_{pil}}$	injected fuel volume quantity of the pilot injection
$\mathbf{q_{pil,tot}}$	total injected fuel volume quantity of the pilot injections
$\mathbf{R^2}$	squared correlation coefficient
RMSE	root mean square error
SOC	start of combustion
SOI	electric start of Injection
$\mathbf{SOI_{main}}$	Electric start of injection of the main pulse
SVM	Support vector machine
t	time
T	temperature
$\mathbf{T_{int}}$	intake manifold temperature
V	volume
VGT	Variable Geometry

	Turbocharger
$\mathbf{X_r}$	EGR rate
WLTP	Worldwide harmonized Light-duty Testing Procedure

Greek symbols

$\gamma = c_p/c_v$	specific heat ratio
λ	relative air-to-fuel ratio
ρ	density
ρ_{SOI}	in-chamber ambient density evaluated at the SOI instant
ρ_{SOC}	in-chamber ambient density evaluated at the SOC instant
τ_{main}	ignition delay of the main pulse
τ_{pil}	ignition delay of the pilot pulse