energies

MDPI

Article

The Enhancement of Machine Learning-Based Engine Models Through the Integration of Analytical Functions

Alessandro Brusa [1,*], Fenil Panalal Shethia [1], Boris Petrone [1], Nicolò Cavina [1], Davide Moro [1], Giovanni Galasso [2] and Ioannis Kitsopanidis [2]

[1] Department of Industrial Engineering, University of Bologna, 40136 Bologna, Italy
[2] Ferrari S.p.A., 41053 Maranello, Italy
* Correspondence: alessandro.brusa6@unibo.it

Abstract: The integration of analytical functions into machine learning-based engine models represents a significant advancement in predictive performance and operational efficiency. This paper focuses on the development of hybrid approaches to model engine combustion and temperature indices and on the synergistic effects of combining traditional analytical methods with modern machine learning techniques (such as artificial neural networks) to enhance the accuracy and robustness of such models. The main innovative contribution of this paper is the integration of analytical functions to improve the extrapolation capabilities of the data-driven models. In this work, it is demonstrated that the integrated models achieve superior predictive accuracy and generalization performance across dynamic engine operating conditions, with respect to purely neural network-based models. Furthermore, the analytical corrective functions force the output of the complete model to follow a physical trend and to assume consistent values also outside the domain of values assumed by the input features during the training procedure of the neural networks. This study highlights the potential of this integrative approach based on the implementation of the effects superposition principle. Such an approach also allows us to solve one of the intrinsic issues of data-driven modeling, without increasing the complexity of the training data's collection and pre-processing.

Keywords: machine learning; neural networks; engine modeling; effect superposition; analytical functions; generalization; fault prediction

check for updates

Citation: Brusa, A.; Shethia, F.P.; Petrone, B.; Cavina, N.; Moro, D.; Galasso, G.; Kitsopanidis, I. The Enhancement of Machine Learning-Based Engine Models Through the Integration of Analytical Functions. *Energies* **2024**, *17*, 5398. https://doi.org/10.3390/en17215398

Academic Editors: Roberta De Robbio and Maria Cristina Cameretti

Received: 24 September 2024
Revised: 25 October 2024
Accepted: 26 October 2024
Published: 30 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The rapid evolution of automotive technology, driven by increasingly stringent emissions regulations and the demand for higher fuel efficiency, has necessitated significant advancements in engine design and calibration. Modern internal combustion engines must meet a complex set of requirements, balancing performance, efficiency, and emissions within a highly constrained operational framework [1]. This balance is particularly challenging due to the diverse conditions under which engines must operate, from low-load urban driving to high-load highway cruising [2,3].

To address these challenges, engine design and calibration processes have evolved to incorporate a variety of advanced techniques and technologies. Among these, the integration of model-based design (MBD) and optimization algorithms stands out as a transformative approach. By leveraging detailed simulations and robust optimization frameworks, engineers can explore a vast design space more effectively than ever before, identifying optimal configurations that would be impractical to evaluate experimentally. Machine learning (ML) techniques are being increasingly introduced into all engine development phases to meet the following objectives.

1. Model-Based Design: Utilizing high-fidelity engine models to simulate performance and emissions under various operating conditions, providing a detailed understanding of engine behavior [4–6]

2. Optimization Algorithms: Implementing multi-objective optimization techniques to identify the best trade-offs between conflicting goals, such as those related to power output, fuel consumption, and pollutant emissions [7,8]
3. Control Strategies: Developing and applying advanced control strategies, including feedback and feedforward mechanisms, to ensure optimal engine performance across the entire operating range [9,10]
4. Data Analytics: Employing real-time data analytics to continuously refine and validate models, ensuring that the optimization process remains aligned with the real-world engine performance [11].

In the field of engine modeling, the integration of analytical functions with ML-based models marks a significant step forward in enhancing the predictive performance and operational efficiency. Traditional engine models, often rooted in complex physical laws and empirical data, face challenges in the identification and calibration of analytical functions, especially when the number of independent variables starts to increase [12]. Conversely, purely data-driven models, particularly those based on artificial neural networks (ANNs), excel in learning complex patterns from data but can struggle with extrapolation beyond their training domain and in maintaining consistency with physical laws [13].

This paper explores the development of hybrid approaches that combine the strengths of traditional analytical methods with modern ML techniques to model engine combustion and temperature indices more effectively [14–16]. By integrating analytical corrective functions into ML-based engine models, we aim to leverage the precision and robustness of physical models alongside the adaptability and learning capabilities of ANNs. This hybridization seeks to address key limitations in the current modeling approaches, such as feature extrapolation, parameter estimation, and the ability to capture physical dynamics [17,18].

The core objective of this study is to demonstrate that integrated models, which blend analytical functions with neural networks, achieve superior predictive accuracy and generalization performance compared to models relying solely on neural networks. The effects superposition principle lies thus at the base of the methodology presented in this work, and this means that each input variable brings a specific contribution to the combustion process. Such a principle has been already implemented in the literature and demonstrated by the authors to include the effects of inert species for the combustion process, like water [19,20]. The innovative contribution of this work is the extension of the methodology previously presented to model complex physical problems, even with a larger number of independent variables. In other words, the main benefit of the proposed approach is the disaggregation of the complex system into simpler problems that depend on a single independent variable, and it can be synthetized as follows:

- The multi-dimensional phenomenon is divided into simpler problems;
- Artificial neural networks are used to capture the effects of variables for which the operating range is well defined (such as the engine speed and load), and it cannot be exceeded during the engine's operation;
- Analytical functions are applied to describe and extend the trend of the output variable (such as combustion, knock, and exhaust gas temperature indices) with respect to those independent variables that could assume values that differ from the calibrated ones; this allows us to increase the extrapolation capabilities of a standard data-driven model;
- The contributions of the artificial neural network and corrective functions are then determined to calculate the final output value.

Through this integrative approach, the enhanced models not only improve the accuracy and robustness of predictions but also maintain consistency with physical trends, even in scenarios where the input features extend beyond the values encountered during training. Such a capability allows us to maintain the good accuracy of the model's predictions, including in the case of anomalous engine operation, and this particular feature makes the models also suitable for failure prediction. Moreover, the improvement does

not impact the effort required for data collection, which relies on standard mapping and calibration activity for a new engine. Additionally, the application of analytical functions for the description of simple physical trends requires fewer engine points for the definition of the calibratable coefficients, with respect to a purely data-driven approach. Indeed, although, for this work, a dedicated dataset has been collected, this is due to the need to also train neural networks and to have a training dataset suitable for a robust comparison between the two methods. Once the hybrid approach is identified as the most robust for this application, a portion of the dataset collected for the engine mapping procedure can be used for model calibration.

The analytical corrective functions play a critical role in ensuring that the output of the complete model adheres to the expected physical behavior, providing a safeguard against unrealistic predictions. This capability is especially crucial when the model is exposed to operating conditions that significantly deviate from the training data, thereby extending the practical applicability and reliability of the engine simulator. The availability of a robust and reliable engine simulator, with the capability to extrapolate the values of the combustion, knock, and exhaust gas temperature indices for values of the input features that differ from the calibrated ones, represents a useful tool for the development of new control strategies [20,21] and the offline adjustment of engine calibrations. Indeed, one of the main targets of such activity is to define a method that can be extended to other indices to develop a complete engine simulator for the prediction of combustion, knock, and exhaust gas temperature indices to achieve the offline design of new driving cycles or to implement new calibration sets.

In the first part of this work, the experimental campaign conducted to collect the data needed for the development of the analytical corrective functions is presented. The description of the complete engine simulator and the indices that have to be modeled are presented, highlighting the types of input variables that can be used to feed the models. The two main approaches considered for the development of the engine models are described and compared in particular for a specific combustion index (50% of the mass fraction burnt, MFB50). Such an index is generally used to develop the methodologies described in this work, and the most accurate approach is then applied to other indices. The methodology utilized to identify the analytical corrective functions used to enhance the performance of the ANNs is described. The accuracy of a purely ANN-based method is compared with that of the hybrid approach when the calibration of the models is performed with steady-state data, while validation is carried out with recordings of driving cycles and dynamic on-track profiles.

This study underscores the potential of hybrid modeling approaches to significantly extend the extrapolation capabilities of engine simulators, offering a more accurate and reliable tool for the prediction of engine performance across a wide range of operating conditions. While it is not discussed in this study, the application of the hybrid approach could also follow a reversed approach, where the main model for the reference conditions is represented by a look-up table (LUT) or an analytical function and the machine learning algorithm works to adjust the main output for dynamic conditions (for example, to include the effect of the sensor dynamics) or to adjust the first value when an additional influencing parameter deviates from the reference conditions. Variables that have a major impact on the combustion process are included as input features of the hybrid model. The engine speed and load, the spark advance, the target lambda value, the temperature of the air at the inlet of the intake valves, and the phases of the intake and exhaust valves are considered as input features, since such variables are also affected by the engine control unit (ECU) strategies that are initiated when particular environmental or vehicle conditions occur. In this way, a wide range of possible conditions that can affect the combustion process are predicted by the model.

In the second part, the development of the complete engine simulator based on the hybrid approach is described, and the results achieved by simulating different types of driving cycles and steady-state conditions are presented. Moreover, some proposals to

further improve the accuracy and robustness of the data-driven models are presented in the last section of this work.

2. Experimental Setup

In this section, the engine used for the experimental campaign and the layout of the testing environment used to control such an engine and to collect data are introduced. Furthermore, a detailed description and critical considerations regarding the experimental campaign conducted are presented.

2.1. Testing Environment

The experimental setup is composed of a high-performance, gasoline-powered internal combustion engine, whose main characteristics are described in Table 1. It is installed in the engine test cell by connecting it to an eddy current passive dynamometer by Borghi & Saveri (Italy).

Table 1. Eight-cylinder engine characteristics.

Displaced volume [cc]	3990
Number of cylinders [#]	8
Stroke [mm]	82
Bore [mm]	88
Connecting rod [mm]	143
Compression ratio [-]	9.5:1
Number of valves per cylinder [#]	4
Combustion system	Spark-ignited gasoline direct injection
Charging system	Single turbocharger for each bank

Each cylinder is equipped with a piezoelectric pressure transducer, in order to collect in-cylinder pressure data, with a sampling frequency of 100 kHz. The pressure sensor used is from Kistler, and its main features are reported in Table 2. Indicating signal conditioning and acquisition is carried out with MASTRO charge amplifiers and the OBI-M2 indicating system provided by Alma Automotive (Italy). The calculations of the MFB50, indicated mean effective pressure (IMEP), and maximum in-cylinder pressure (Pmax) combustion indices use low-pass-filtered signals, with a cut-off frequency of 3 kHz, and a windowed in-cylinder pressure trace, while the same signal is band-pass-filtered in order to calculate the maximum amplitude pressure oscillation (MAPO) index for the knock intensity.

Table 2. In-cylinder pressure sensor characteristics.

Pressure Range [bar]	**0 to 250**
Overload [bar]	300
Sensitivity [pC/bar]	−37.0
Natural frequency [kHz]	>215

The main temperatures of the system (air temperature inside intake manifold, exhaust gas temperature in exhaust runners and at turbine inlet section) are acquired with thermocouples connected to the test bench management system, whose characteristics are summarized in Table 3.

Table 3. Thermocouple characteristics.

Type	**K**
Diameter [mm]	3.3
Accuracy [°C]	±2.2

The thermocouple measurement chain is composed of a National Instruments (Austin, TX, USA) Compact-Rio 9024, with module 9213, which has a sample frequency of 100 Hz. The actual value of the temperature measured for each stationary engine point and condition is acquired after the end of the transient phase, to obtain the characteristic temperature at a given working condition.

The fuel used to feed the engine is gasoline with research octane number (RON) 98 for the whole experimental procedure, as it is not required that the engine model is sensitive to different fuel properties. Furthermore, all data available from the industrial partner of this project that can be used for validation, testing, and analysis are collected using the same type of gasoline. The thermodynamic conditions of the air inside the intake manifold are controlled thanks to the turbocharging system, for the air pressure, and to the water-to-water heat exchange system of the intercooler for the air temperature.

As seen in Figure 1, an engine control unit (ECU) for each bank controls the engine, equipped with ETAS hardware and software tools to manage different working conditions and impose the actuations required by each specific test. Furthermore, real-time communication between the ETAS tools and the test bench acquisition system allows the collection of iso-frequency and phased data between the actuations and the data sensed by the engine test bench.

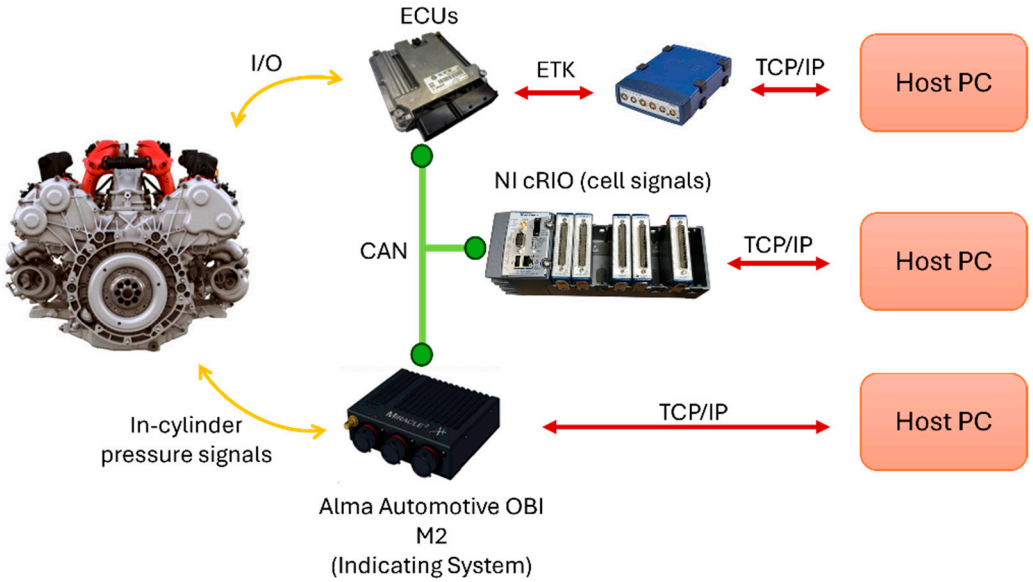


Figure 1. Functional layout of communication loop.

2.2. Experimental Campaign

To highlight the advantages of the hybrid approach followed by the authors, and the reasons for which this path is chosen, in the following sections, a comparison will be presented between this one and a pure ANN methodology. To perform a consistent comparison, a dataset obtained from standard calibration activities is used to train the purely ANN-based model (1200 points are used for this task), given its ability to manage a large amount of data. A portion of this dataset can be also used to perform model validation and testing. Meanwhile, to calibrate the whole hybrid model, specific custom-made tests are required to ensure a robust comparison between the two approaches, as mentioned earlier.

The experimental tests conducted have the main goal of building a consistent dataset that is able to successfully train the ANN and calibrate the analytical corrective functions.

The neural network must be capable of modeling the value of the MFB50 at different engine working points, with standard calibrations of the main actuations and reference inlet air conditions. Thus, it must be sensitive to the variations in the combustion phasing with different engine speeds and loads. For this purpose, a wide grid of engine operating points is collected to avoid the risk of extrapolation during the model's inference. This goal

is fulfilled by testing the engine up to its physical minimum and maximum values for both the engine speed and load, avoiding the possibility of simulating maneuvers or operating conditions outside of these boundaries. As visible in Figure 2, which shows the steady-state data recorded for the ANN training, the operating points of the grid are quite narrow to ensure good interpolation capabilities, avoiding being too well fitted, which could lead to a risk of overfitting (a total amount of 89 engine points are collected). The engine speed and load are normalized with respect to the maximum value and then converted to percentages. This normalization process for the engine speed and load is followed throughout this paper.

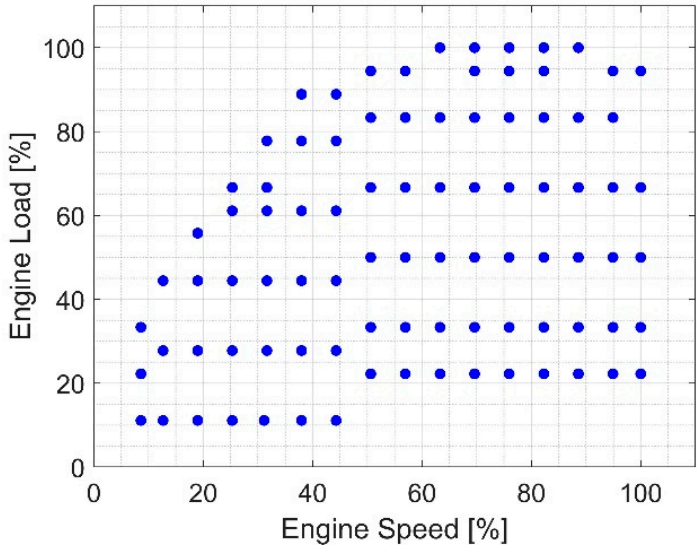


Figure 2. Engine points of data collected for ANN training.

Furthermore, different testing activities are conducted to collect a dataset that is useful to calibrate the analytical corrective functions. Keeping in mind that the core idea of the model is to leverage the use of the superposition principle, each test consists of a sweep of each independent variable that acts as input for the model: the spark advance (SA), lambda (λ), variable valve timing (VVT), and intake manifold temperature (Tans). Considering all these elements, fewer engine points are considered necessary with respect to the first dataset, as seen in Figure 3. The selection of these points is based on the goal of covering the whole engine speed and load map, while isolating the effect of each variable through their sweeping. In particular, the engine map can be divided into four zones to be explored: a low load and low speed, a high load and low speed, a low load and high speed, and a high load and high speed. These cover both aspirated and boosted zones of the engine's working areas, from low to high speeds. As another challenge in the current work, a small number of engine operating points are chosen, ensuring that a reasonable number of total points is tested, considering that, for each engine point, a certain number of sweep tests is performed. To ensure the robustness of the corrective functions (a total of 615 points are collected), 19 engine points are used for each number of sweeps performed for each engine point and independent variable.

An important characteristic of these tests is that, during the variation in each input, all other independent variables are kept as constant as possible, as visible in Figures 4 and 5. If it is an actuated variable, it is fixed at the same value as for standard calibration; meanwhile, if it is a sensed parameter, it must be close to the reference value. This element is very important to isolate each input's effect on the combustion phasing. The numerical values of the variables are omitted for confidentiality; thus, it is important to highlight that the maximum value of oscillation accepted is ±5%, especially for those parameters that are difficult to control (e.g., the intake manifold temperature).

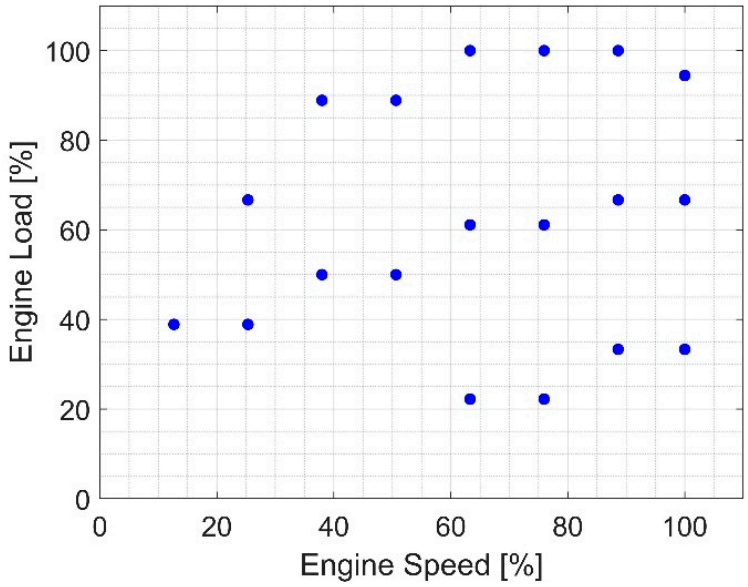


Figure 3. Engine points of data collected for corrective function calibration.

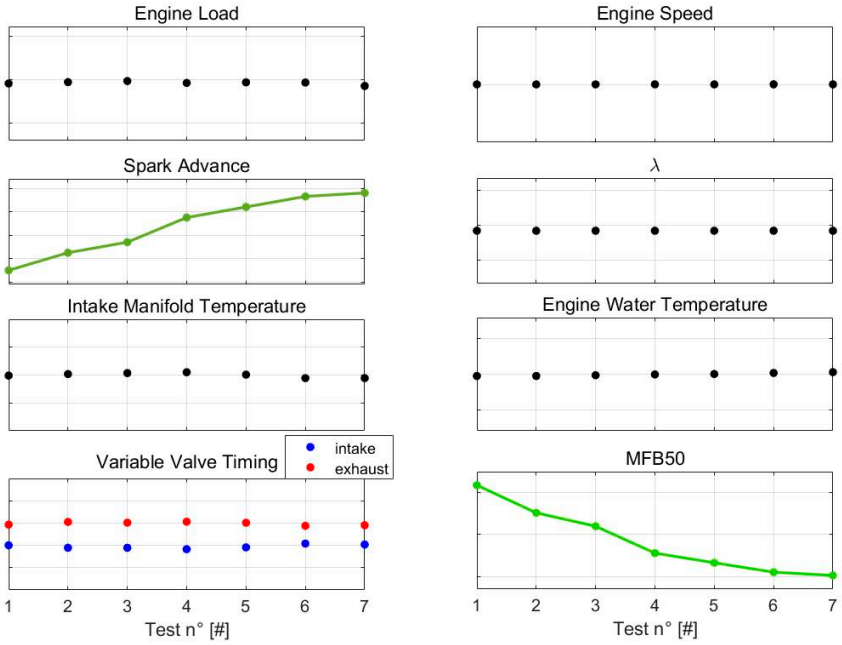


Figure 4. Spark sweep effect on MFB50 isolated for a specific engine point.

Additionally, some tests are carried out to build a proper validation dataset for the model: both steady-state engine points with a simultaneous change in multiple inputs, in order to stress the combined effect of all variables involved in the combustion process, and transient on-road profiles replicated at the engine test bench from standard homologation cycles to track their profiles.

Figure 6 shows an example of a standard homologation cycle on the left and a track test profile on the right side, represented both by the normalized driver pedal position and normalized engine speed. It is also important to highlight that, throughout the whole validation process, transient tests with both standard calibration and different control strategies are performed. This is crucial to fully explore the capability of the developed model to reproduce combustion metrics in transient conditions and non-conventional control srategies.

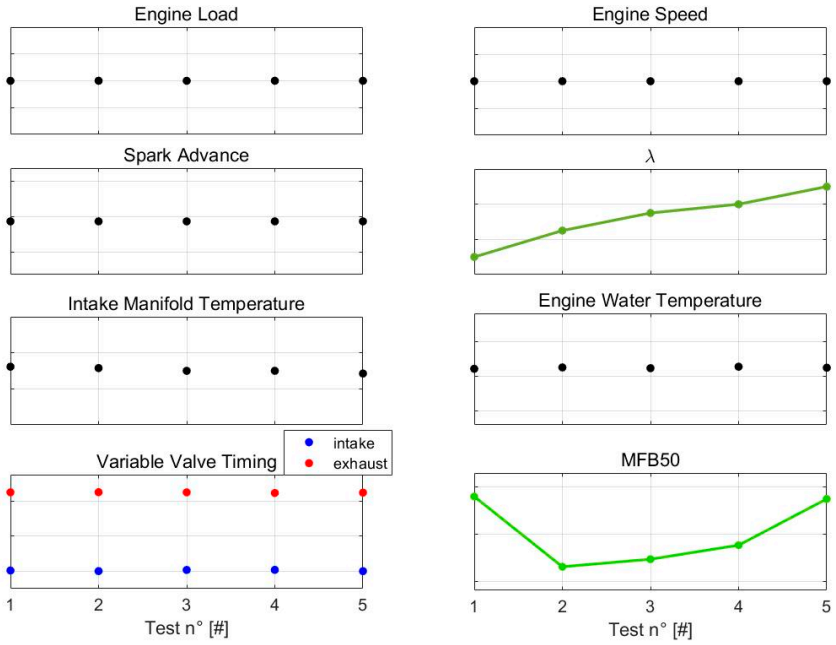


Figure 5. Lambda sweep effect on MFB50 isolated for a specific engine point.

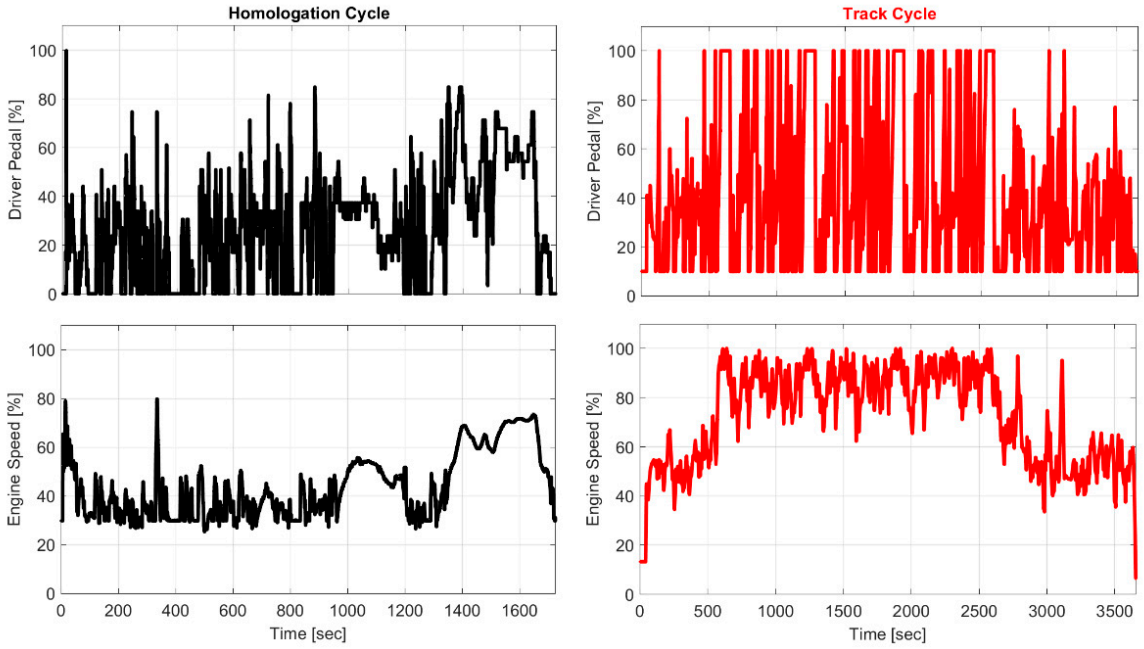


Figure 6. Transient engine profiles.

3. Comparison and Description of Approaches

The two approaches compared in this work are as follows: the first relies solely on pure ANN models, while the second, referred to as the hybrid approach, combines ANNs with analytical corrective functions to account for the effects of various variables and actuations. The accuracy of both approaches is compared by evaluating the error under steady-state conditions, and the hybrid one is demonstrated to be the most accurate, especially due to its stronger extrapolation capabilities. The complete research workflow is reported in Figure 7.

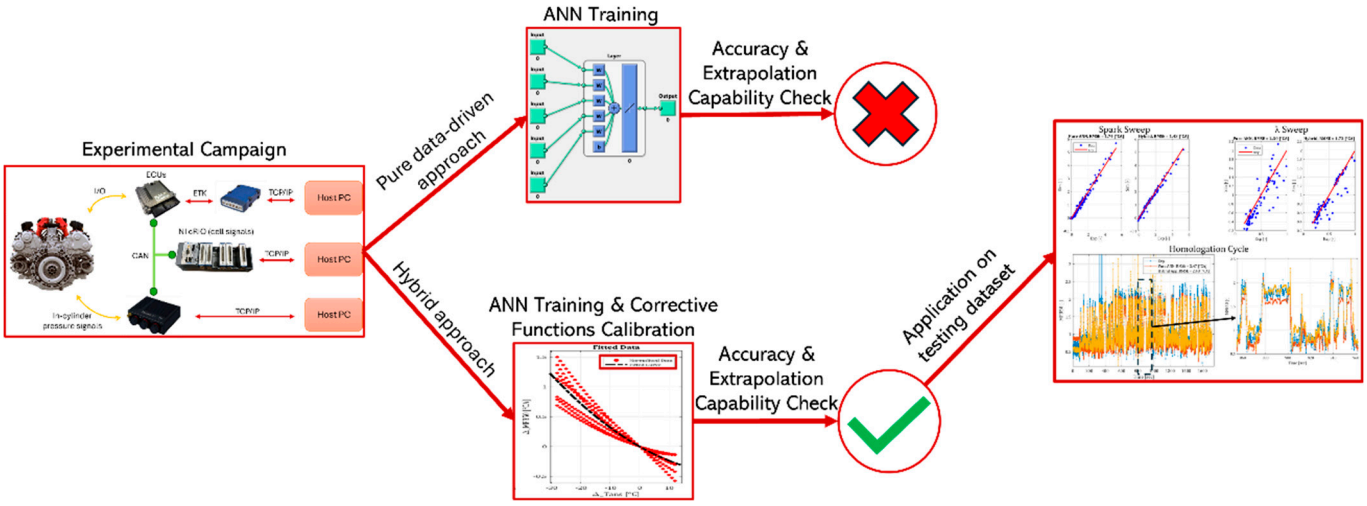


Figure 7. Research flowchart.

3.1. Pure ANN Approach

The purely ANN-based approach has been developed in our previous work [22], in which an automatic calibration algorithm was developed to calibrate/train the ANN for the given index. This algorithm uses both steady-state and transient data to find the optimal configuration of the network. Sufficiently accurate results have been obtained in different steady-state and transient tests. However, as is well known, these ANN models do not have a strong extrapolation capacity and, in some cases, could lead to an unphysical result. For example, the physical relationship between SA and MFB50 is well known, i.e., for an anticipated SA, the resulting MFB50 is also anticipated and vice versa. An ANN, being a black box model, does not learn this physical relationship, and, when the inputs being tested are outside the dataset used for training (in the extrapolation zone), the model for a retarded SA predicts an anticipated MFB50, leading to an unphysical result. This behavior can be seen in Figure 8. In the left-hand plot, the pure ANN model predicts a very retarded MFB50 (MFB50 = 10–12), whereas the experimental value is much more anticipated (MFB50 = 7–8). Furthermore, as shown in the right-hand plot, during a dynamic test, the variable valve timing changes at a certain time (due to certain ECU strategies), and the pure ANN model (red line) predicts an anticipated MFB50 instead of a retarded one.

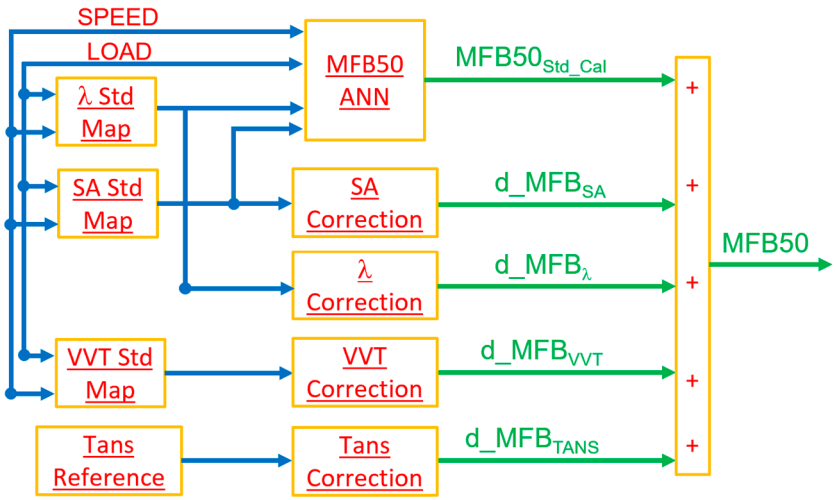


Figure 8. MFB50 calculation block chain.

3.2. Hybrid Approach

The need to extrapolate is one of the driving factors in the development of the second approach, i.e., the hybrid approach. This approach tries to exploit the superposition principle by combining the ANN approach with analytical corrective functions. To explain the methodology, the combustion index, the MFB50, is used, and a similar approach is applied to the other indices. This approach is composed of two parts, i.e., the first part is an ANN, and the second part consists of the analytical corrective functions.

The ANN is trained with steady-state data across a wide engine operating range, which are recorded with the reference condition of Tans and the standard/mapped calibrations for all actuations, such as the SA, λ, and VVT. The inputs to this ANN are the engine speed and load, the SA, and the λ. This model is trained by using 100% training data, because its sole purpose is to predict the MFB50 under standard calibrations. The calibration of the ANN is performed using the same, previously developed algorithm [22], with a slight modification due to using 100% training data.

For the analytical corrective functions, various sweep tests, as mentioned in the previous sections, are carried out. An individual corrective function for each of these actuations is developed (as explained in the model development section, i.e., Section 4), which captures the effect of the variation in the given actuation. At the end, all of the corrective functions, along with the output of the ANN, are summed to give the final calculated MFB50. Figure 8 shows the general block scheme of this calculation chain.

Figure 9 shows some results to demonstrate that the hybrid approach has superior extrapolation capabilities with respect to the purely ANN-based approach and why it is chosen as the final approach. The graph on the left in Figure 9 illustrates a steady-state test, named Test 1, where multiple actuations have been changed at the same time, i.e., the SA, injection, VVT, etc. Meanwhile, the graph on the right refers to a dynamic test where there was a change in the VVT (both intake and exhaust) following certain strategies of the ECU. The results are compared using the root mean square error (RMSE) index. As seen from the results, the hybrid approach captures these changes with higher accuracy with respect to the ANN-based model and can also follow the physical trend of the phenomenon. Following these results, the hybrid approach is selected for subsequent analysis. The values of the MFB50 are normalized and one unit equals 10 °CA; this criterion for the normalization of this index is implemented throughout this paper. The Y axis of the VVT graph is obscured for confidentiality reasons.

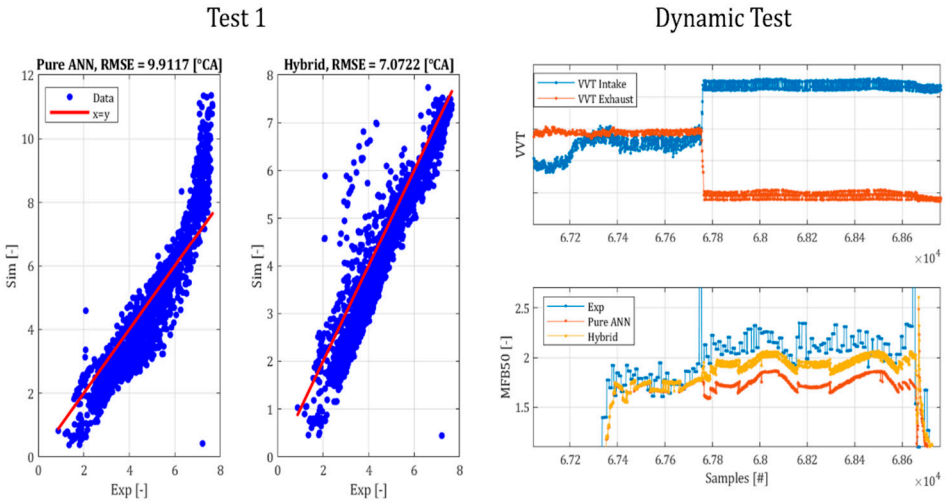


Figure 9. MFB50 results comparison between pure ANN and hybrid approaches.

4. Models' Development

As mentioned earlier, the hybrid approach is composed of two parts. The first part is composed of an ANN that has been trained on the data shown in Figure 2. These data

actuate the standard calibrations for various variables and at the reference Tans. The second part consists of the analytical corrective functions developed from the data shown in Figure 3. It must be highlighted that all models developed in this work refer to the mean cylinder.

4.1. MFB50 Model

4.1.1. ANN

The ANN for the MFB50 is trained using 100% of the steady-state data recorded, as shown in Figure 2, and the inputs to this model are the engine speed, the load, the SA, and the λ. Since this model must predict the value of the MFB50 only under the reference conditions and the standard calibrations, 100% training data are used to avoid extrapolation problems in the ANN. The calibration of the ANN itself is achieved with the self-developed algorithm [22], with a slight modification. The final configuration of this ANN is as shown in Table 4.

Table 4. MFB50 ANN configuration.

Training algorithm	Bayesian regularization
Activation function	Logsig
Number of neurons	15
Number of epochs	3000
Training dataset size	89 engine points

4.1.2. SA Correction

As is well known from the literature, the relationship between the SA and the MFB50 is characterized by a parabolic trend, which can be described by a second-order polynomial [21]. Taking advantage of this relationship, the SA corrective function has been developed, where the function outputs the ΔMFB50, which is a function of the ΔSA. Starting from the spark sweep data shown in Figure 3, the MFB50 and the SA axis are normalized following Equations (1) and (2).

$$\Delta SA = SA_{Actuated} - SA_{StdMap} \tag{1}$$

$$\Delta MFB50 = MFB50_{SAActuated} - MFB50_{SAStdMap} \tag{2}$$

As seen in Figure 10, the normalized data collapse well into a single curve representing the entire operating range of the engine. A single curve, a second-order polynomial, is fitted over these normalized data. A single curve representing the entire engine domain also allows us to simplify the calculation chain, without the need to have further maps/look-up tables, etc. For the coefficients, this reduces the error in the calculation. This relationship is characterized by Equation (3):

$$\Delta MFB50 = A \cdot \Delta SA^2 + B \cdot \Delta SA + C \tag{3}$$

where A, B, and C are the equation parameters.

4.1.3. λ Correction

Starting from the experimental data of the λ sweeps, it has been observed that λ has a parabolic relationship with the MFB50, as shown in Figure 11. First, the experimental data are fitted with a second-order polynomial for each engine point test. This is carried out to find the value that λ has when the minimum MFB50 has been reached for the given engine load and speed, and a map of this minimum value is created as a function of the engine speed and load. This is performed because, while developing the corrective function, the data have been normalized with respect to the λ value that has the minimum MFB50, instead of the standard mapped value of λ. Furthermore, due to the normalization performed with respect to the minimum value, there is a need to introduce this corrective

function twice in an algebraic subtraction manner to avoid erroneous corrections being applied, especially when the standard mapped calibrations are actuated. Then, these data, fitted with a second-order polynomial, are normalized using Equations (4) and (5).

$$\Delta\lambda = \lambda_{range} - \lambda_{min} \tag{4}$$

$$\Delta MFB50 = MFB50_{\lambda range} - MFB50_{min} \tag{5}$$

where

- λ_{range} = the range of λ that covers values between rich and slightly lean;
- λ_{min} = the value where the MFB50 is the minimum for the given engine point.

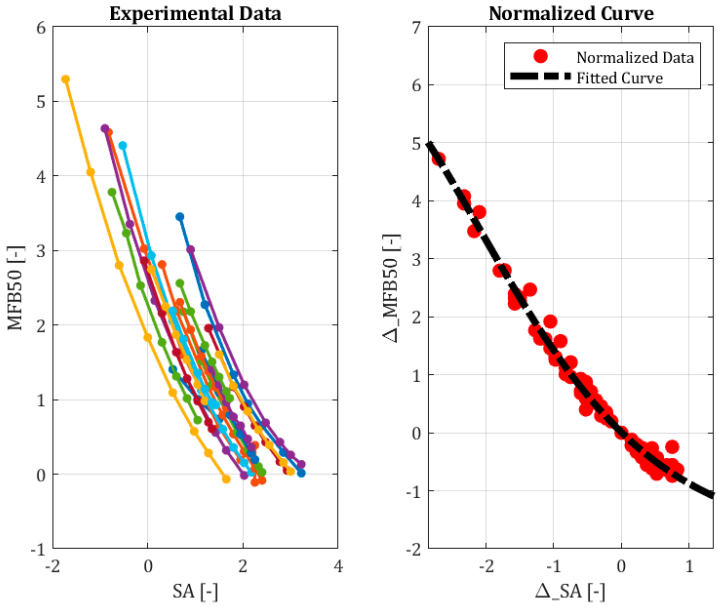


Figure 10. SA sweep and corrective function for MFB50. Each colored line in the left graph represents an engine point (as a couple of the engine speed and engine load).

Once the data have been normalized, a single curve, namely a second-order polynomial as given by Equation (6), is fitted, which represents the relationship of $\Delta MFB50$ with respect to Δλ over the entire operating range of the engine, as shown in Figure 10.

$$\Delta MFB50 = D{\cdot}\Delta\lambda^2 + E{\cdot}\Delta\lambda + F \tag{6}$$

where D, E, and F are the equation parameters.

4.1.4. VVT Correction

As is well known, variable valve timing systems are being increasingly used in modern engines as they help with emissions and performance. These essentially affect the flow through the intake and exhaust valves, thereby affecting the combustion itself. Therefore, it is necessary to incorporate this effect when evaluating the MFB50. As seen from the VVT sweeps in Figure 12, the data have similar behavior to λ; hence, a similar approach is followed to develop the corrective function for the VVT. This figure shows the results for the intake side; similar curves have been obtained for the exhaust side. The equations used to normalize the data and fit a second-order polynomial are as shown in Equations (7)–(9).

$$\Delta VVT = VVT_{range} - VVT_{min} \tag{7}$$

$$\Delta MFB50 = MFB50_{VVTrange} - MFB50_{MFB50min} \tag{8}$$

$$\Delta MFB50 = G{\cdot}\Delta VVT^2 + H{\cdot}\Delta VVT + I \tag{9}$$

where

- VVT_{range} = the range of the VVT that covers values between anticipating and retarding VVTs;
- VVT_{min} = the value where the MFB50 is the minimum for the given engine point;
- G, H, and I are the equation parameters.

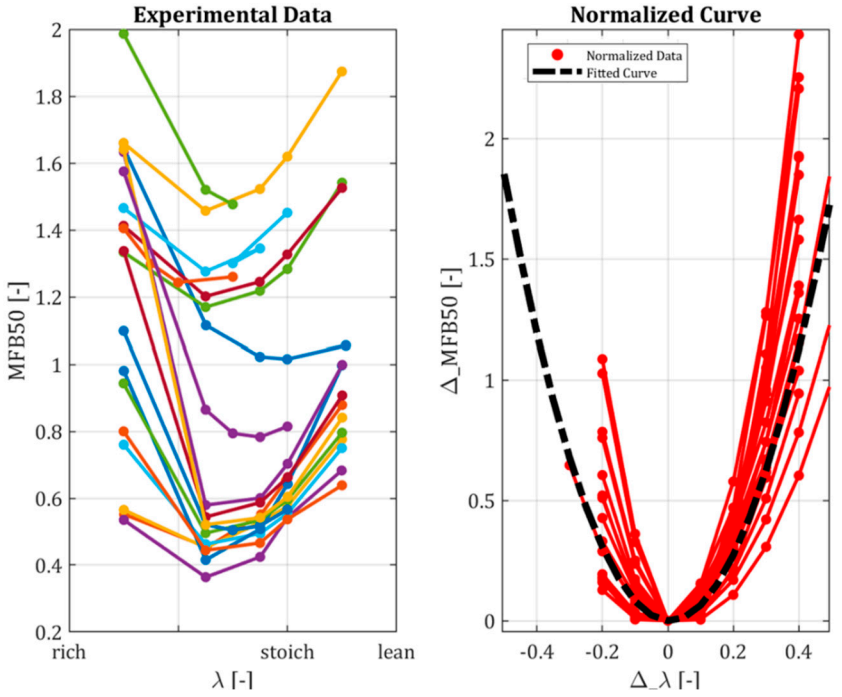


Figure 11. Lambda sweep and corrective function for MFB50. Each colored line in the left graph represents an engine point (as a couple of the engine speed and engine load).

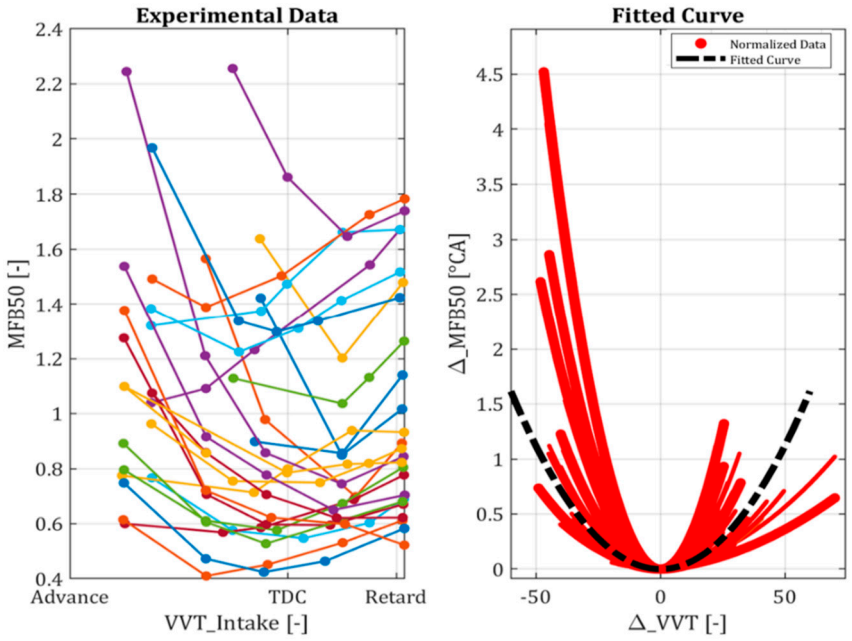


Figure 12. VVT sweep and corrective function for MFB50. Each colored line in the left graph represents an engine point (as a couple of the engine speed and engine load).

4.1.5. Tans Correction

For a given engine speed and load, SA, λ, and VVT, the intake manifold air temperature affects the combustion, i.e., if this temperature is higher, it anticipates the combustion, and, if it is lower, it retards the combustion. A corrective function has been developed to capture this effect, and the experimental data have been normalized following Equations (10) and (11). A single curve is fitted over these points, as shown in Figure 13, which is a second-order polynomial given by Equations (10)–(12).

$$\Delta Tans = Tans_{Actuated} - Tans_{Reference} \tag{10}$$

$$\Delta MFB50 = MFB50_{TansActuated} - MFB50_{TansReference} \tag{11}$$

$$\Delta MFB50 = J \cdot \Delta Tans^2 + K \cdot \Delta Tans + L \tag{12}$$

where D, E, and F are the equation parameters.

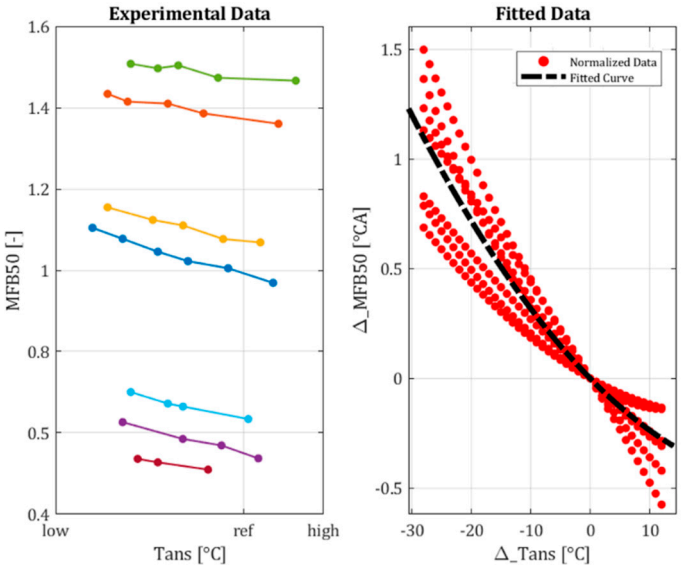


Figure 13. Tans sweep and corrective function for MFB50. Each colored line in the left graph represents an engine point (as a couple of the engine speed and engine load).

The complete and detailed calculation block chain for the MFB50 model is shown in Figure 14. In the figure, the VVT block is shown once, but it represents the calculation for both the intake and exhaust sides.

4.2. Pmax Model

The Pmax model is based on a 2D LUT that describes the trend of the maximum in-cylinder pressure as a function of the engine load and the MFB50. Since the Pmax can be described using only two inputs, and to obtain a robust model that is able to capture the physical trend of the index, it is decided to use a look-up table instead of an ANN.

As is well known, for a highly retarded combustion, i.e., in the expansion phase, the maximum pressure value becomes independent of the MFB50, and it corresponds to the maximum pressure in the motoring condition. This is taken into consideration when fitting the data and then estimating the Pmax through the model.

The fitting function is selected to evaluate the model's performance in terms of the RMSE and R^2. The best results are found when the data are fitted with a "thin-plate spline" function. To avoid the unphysical calculation of the Pmax when the combustion is

very retarded, the Pmax value chosen is the motoring Pmax, which can be estimated by considering the process as polytropic compression, using Equation (13):

$$pV^n = Const \tag{13}$$

where

- p = pressure;
- V = volume;
- n = polytropic index.

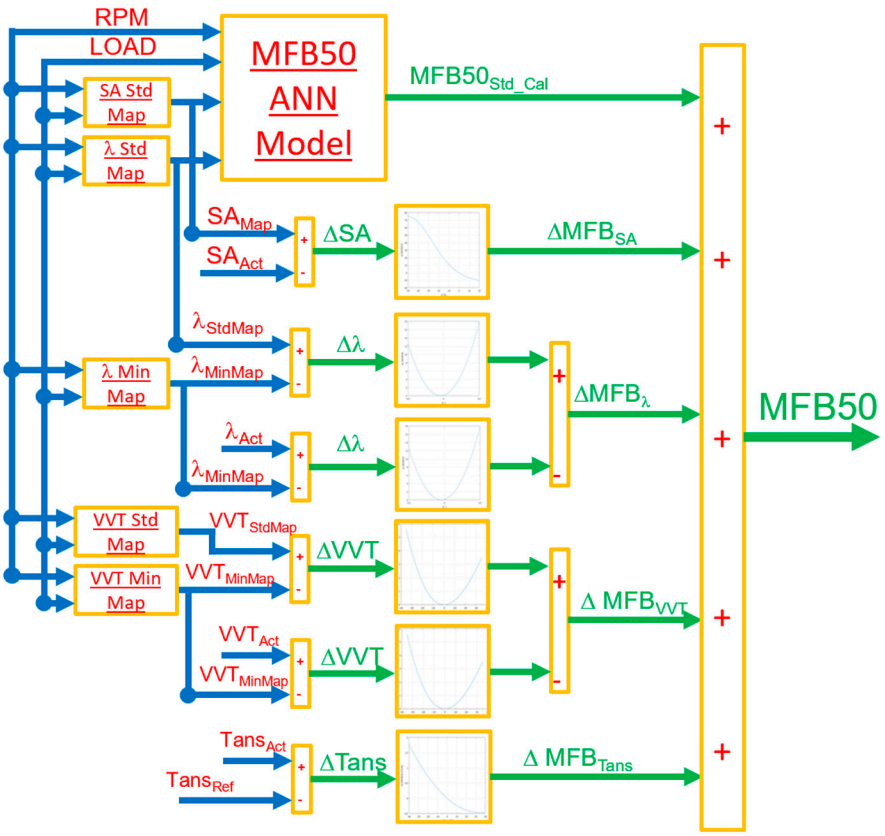


Figure 14. Complete and detailed calculation chain for evaluation of MFB50.

The Pmax is therefore evaluated using the engine load and the MFB50 derived from the MFB50 model. Figure 15 shows the calculation block chain and the final fitted surface of the Pmax as a function of the engine load and MFB50. It must be highlighted that the MFB50 model shown in Figure 15 is the complete model as explained in the previous section; here, it is simplified due to space constraints. The values of the Pmax have been normalized with respect to the maximum and then converted into percentages, and the criterion for the normalization of this index is applied throughout this paper.

4.3. IMEP Model

The approach used to evaluate the IMEP is similar to that for the Pmax. The model is again a 2D LUT as a function of the engine load and the MFB50. The experimental data are fitted with the *poly22* function, which allows us to maintain the well-known bell-shaped efficiency curve and has high accuracy. Equation (14) represents the IMEP analytical model.

$$IMEP = p_{00} + p_{10}{\cdot}Load + p_{01}{\cdot}MFB50 + p_{20}{\cdot}Load^2 + p_{11}{\cdot}Load{\cdot}MFB50 + p_{20}{\cdot}MFB50^2 \tag{14}$$

where

- *Load* = engine load;
- p_{xy} = coefficients of the *poly22* function.

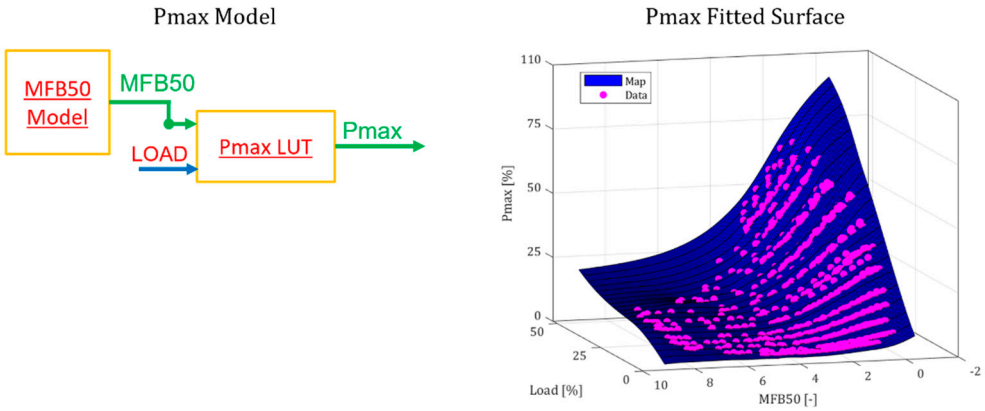


Figure 15. Pmax model block scheme and fitted surface.

Figure 16 shows the calculation block chain and the fitted surface for the IMEP. The values of the IMEP have been normalized with respect to the maximum and then converted into percentages, and the criterion for the normalization of this index is applied throughout this paper.

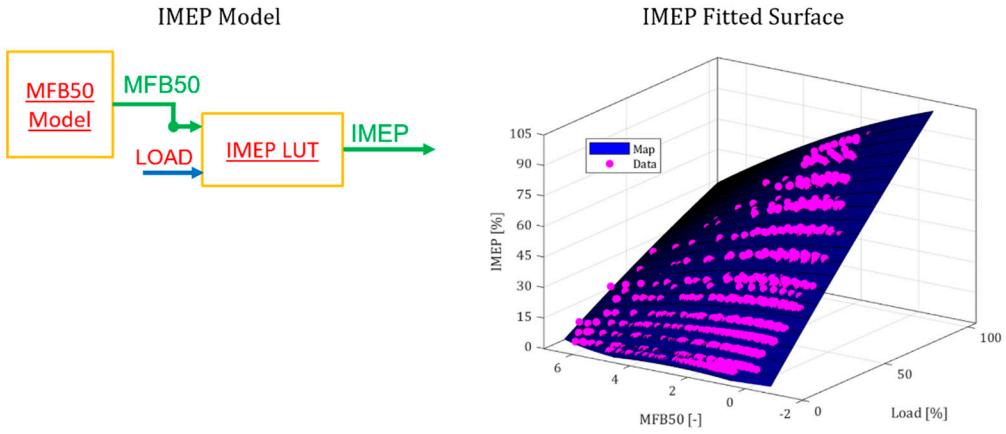


Figure 16. IMEP model block scheme and fitted surface.

4.4. Knock Model

Knock is widely recognized as a stochastic phenomenon, and its behavior is most effectively analyzed through statistical methods. For a given engine speed, load, SA, and λ, the cyclic MAPO statistical distribution can be modeled using either a gamma or a log-normal distribution probability density function (PDF). However, the log-normal PDF shares key characteristics with the Gaussian distribution, enabling the characterization of the distribution with just two parameters, the mean value μ and the standard deviation σ, which can be computed using the following equations [23]:

$$\mu = \log(MAPO50) \tag{15}$$

$$\sigma = \frac{\log(MAPO98.5) - \mu}{2.1725} \tag{16}$$

where

- *MAPO50* corresponds to the 50th percentile of MAPO;
- *MAPO98.5* represents the 98.5th percentile of MAPO;
- μ is the mean value;

- σ is the standard deviation.

Following Equations (15) and (16), two separate ANNs have been trained to model μ and σ following the algorithm developed previously [22]. The inputs to these models are the Pmax as calculated with the Pmax model, the MFB50 as calculated with the MFB50 model, the engine speed and load, and the λ. In total, 70% of the data, as represented in Figures 2 and 3, is used for the training of these two ANNs. The ANN approach is used for the knock model because it takes as input the calculated MFB50 (calculated by the hybrid approach) and the calculated Pmax, which consider the effects of various actuations/variables. The configurations of both of these ANNs are shown in Table 5, and the calculation block chain is shown in Figure 17.

Table 5. The μ and σ ANN configuration.

	μ	σ
Training algorithm	Bayesian regularization	Bayesian regularization
Activation function	Elliotsig	Elliotsig
Number of neurons	20	26
Number of epochs	3000	3000
Training data	564 engine points	564 engine points

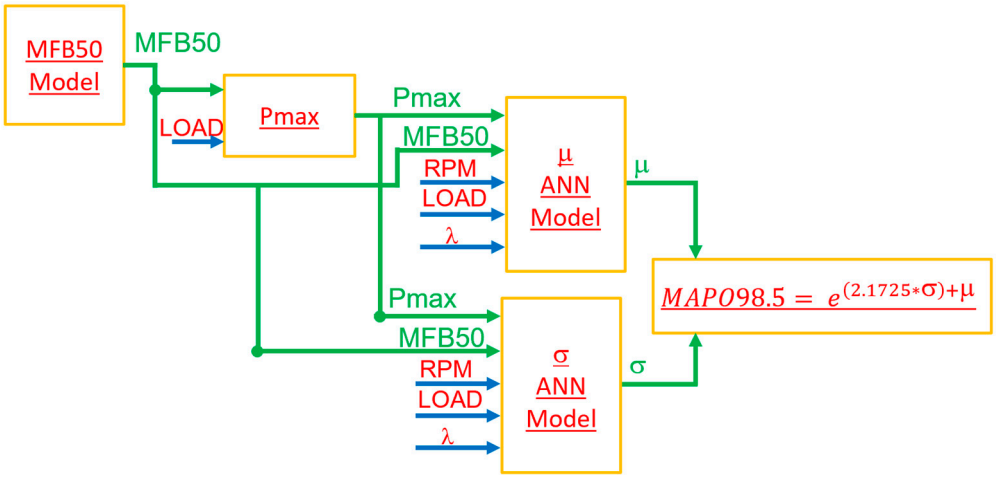


Figure 17. MAPO model block scheme.

4.5. Exhaust Gas Temperature Model

The exhaust gas temperature model, measured immediately after the exhaust valves, has been developed following a similar approach to that used for the MFB50, i.e., following the hybrid approach with individual corrective functions for all of the variables/actuations. The values of the exhaust gas temperature have been normalized with respect to the maximum value and converted to percentages, and this method of normalization for this index is applied throughout this paper.

4.5.1. SA Correction

The SA correction for the exhaust gas temperature has been developed similarly to the MFB50 SA correction, with the exception that the corrective function is a 2D LUT instead of a 1D table, as for the MFB50. This 2D LUT outputs the ΔT_{Exh} (exhaust gas temperature), which is a function of ΔSA and the engine load. To develop the corrective function, the experimental data have been normalized following Equations (17) and (18):

$$\Delta SA = SA_{Actuated} - SA_{StdMap} \quad (17)$$

$$\Delta T_{Exh} = T_{Exh_{SAActuated}} - T_{Exh_{SAStdMap}} \quad (18)$$

Once the data have been normalized, a surface is fitted using the 'poly12' function, which gives an R^2 of 0.97. The equation for this polynomial is given in Equation (19). Figure 18 shows the experimental data and the map of the corrective function. The value of the exhaust gas temperature has been normalized with respect to the maximum and converted to a percentage value. This criterion is applied throughout this paper for this index.

$$\Delta T_{Exh} = p_{00} + p_{10}\cdot Load + p_{01}\cdot \Delta SA + p_{11}\cdot Load\cdot \Delta SA + p_{02}\cdot \Delta SA^2 \quad (19)$$

where

- $Load$ = engine load;
- p_{xy} = coefficients of *poly12* function.

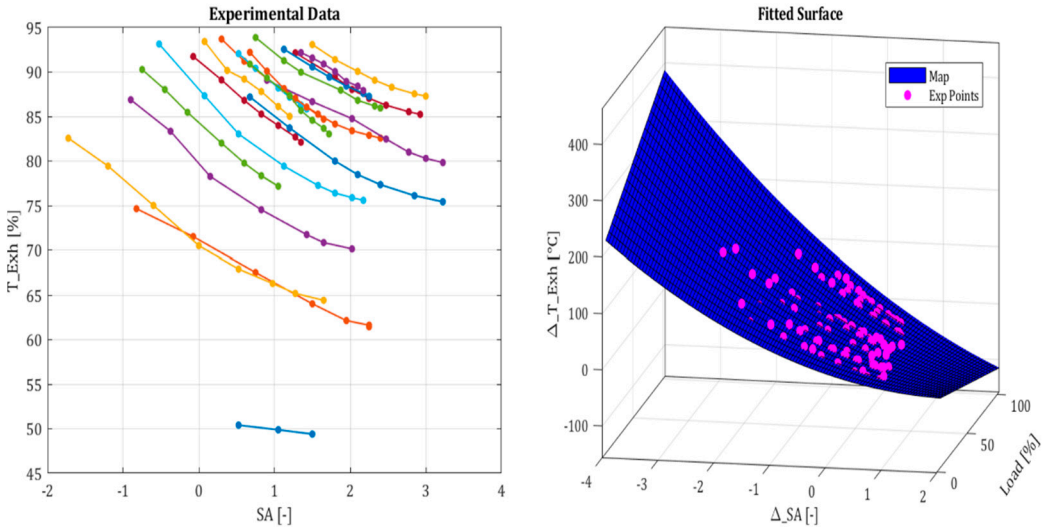


Figure 18. SA sweep and corrective function for exhaust gas temperature. Each colored line in the left graph represents a engine point (as a couple of the engine speed and engine load).

4.5.2. λ Correction

A slightly different approach has been followed for the λ correction. As seen in Figure 19 (an example of one engine point), the experimental data show that the λ trend for the exhaust gas temperature cannot be captured by a single curve/function. For the rich zone, the line has a different slope with respect to the lean zone. Therefore, two separate functions have been developed, i.e., one for the zone between the rich and stoichiometric zones and another for the stoichiometric and lean zones of λ. The function is a 2D LUT as a function of the engine load and λ. The value of the exhaust gas temperature is the maximum when λ is stoichiometric; hence, this is used to normalize the data to develop the corrective function, given by Equations (20) and (21).

$$\Delta\lambda = \lambda_{range} - \lambda_{stoich} \quad (20)$$

$$\Delta T_{Exh} = T_{Exh_{\lambda range}} - T_{Exh_{\lambda Stoic}} \quad (21)$$

where

- λ_{range} = the range of λ that covers values between rich and slightly lean;
- λ_{stoic} = the stoichiometric value of λ;
- $T_{Exh\lambda range}$ = the exhaust temperature for the tested value of λ;
- $T_{Exh\lambda stoic}$ = the exhaust temperature for the stoichiometric value of λ.

Once the data have been normalized, a surface is fitted over these data with a 'thin-plate spline' function for each of the two zones, as shown in Figure 20.

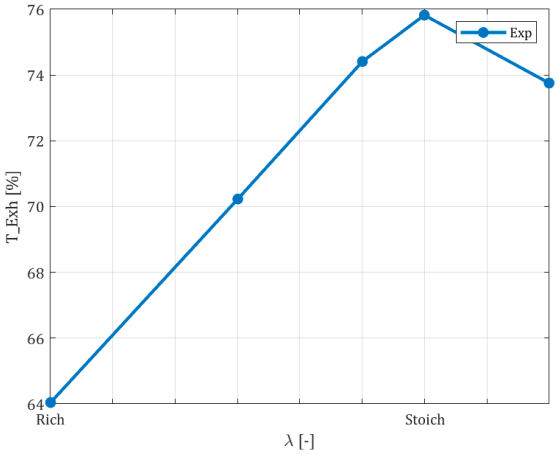


Figure 19. Lambda sweep experimental data for exhaust gas temperature for one engine point.

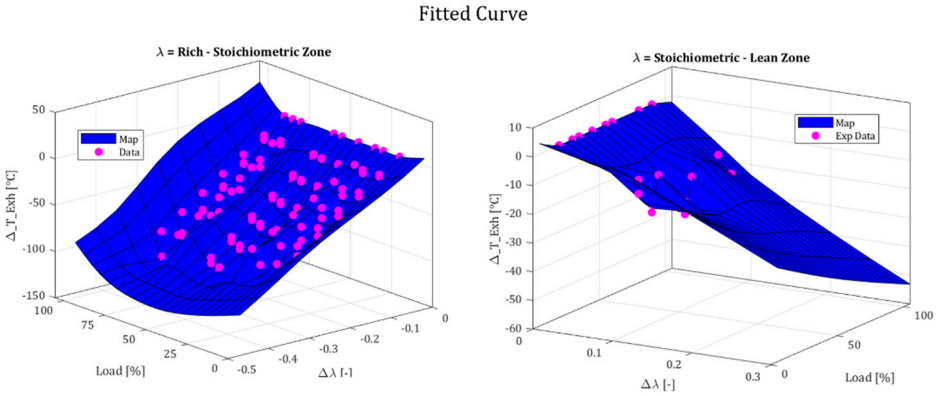


Figure 20. Lambda corrective function for exhaust gas temperature.

4.5.3. VVT and Tans Correction

It can be observed from the experimental data shown in Figure 21 that, for a given engine point (with each curve representing the coupling of the engine speed and load), there is a negligible effect (less than 2%) of the VVT and Tans on the exhaust gas temperature. Hence, the effects of these actuations are eliminated from the calculation chain.

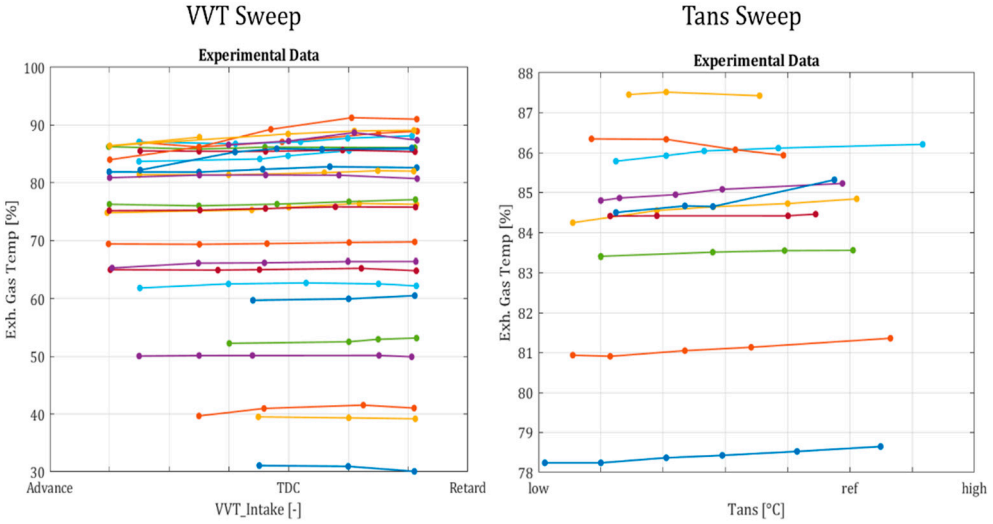


Figure 21. VVT and Tans sweep experimental data. Each colored line in the left graph represents a engine point (as a couple of the engine speed and engine load).

Figure 22 shows the complete calculation chain for the evaluation of the exhaust gas temperature.

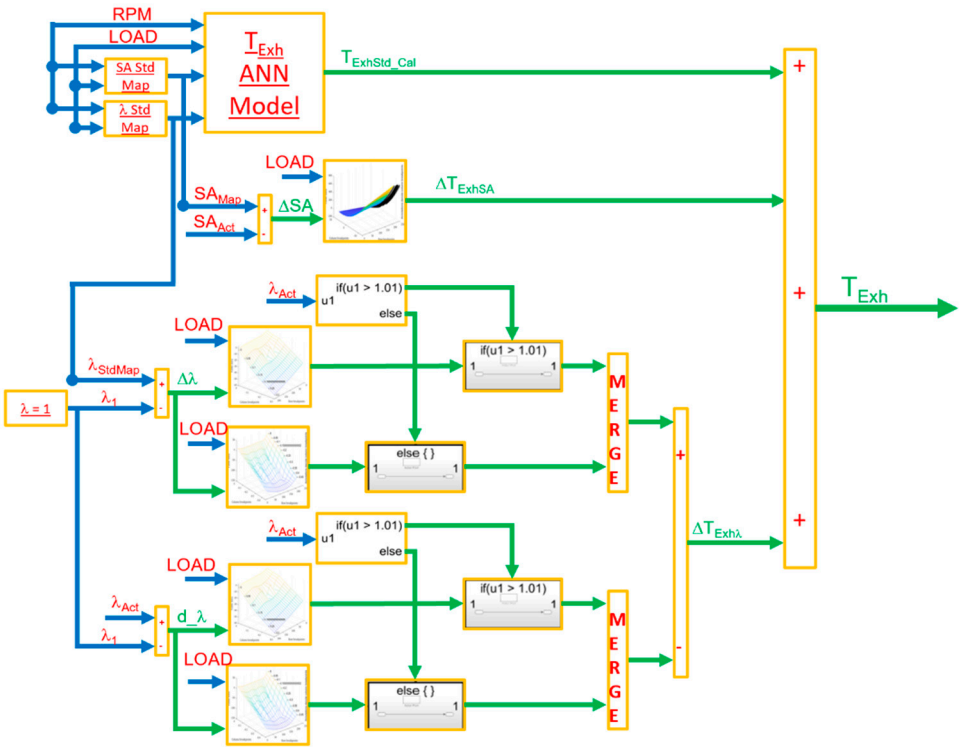


Figure 22. Exhaust gas temperature calculation block chain.

4.5.4. Thermocouple Model

The developed exhaust gas temperature model described above was developed and validated using steady-state data. Nevertheless, to validate the model in transient conditions, it is necessary to model the behavior of the thermocouple itself. The model for the thermocouple has been developed in our previous work [9]. It works as a first-order system with a time constant τ. This τ has been calibrated using a self-developed optimization algorithm that outputs the optimal value τ of under different driving cycles/transient tests.

5. Simulations and Results

All of the models developed above have been validated under steady-state conditions and then under transient conditions. The simulations were performed in the Simulink environment, and the results are compared for the hybrid and purely ANN-based approaches.

Firstly, the models were validated using sweep tests, as shown in Figure 3. For further validation and to ensure the generalization of the models, they were tested on other steady-state data and transient data that were not used during model calibration. 'Test 1' in this section refers to steady-state tests that included multiple variables/actuations changing at the same time, following certain ECU strategies. In the following, figures of the results are presented, for synthesis reasons, for some tests only, but the numerical results are shown for all validation tests.

5.1. MFB50

Table 6 shows the results achieved for the MFB50 in various steady-state and dynamic tests. As seen from the results, the hybrid approach has similar/higher accuracy with respect to the purely ANN-based approach. It must be highlighted that, even if the results in terms of the RMSE between the two approaches, as seen in Table 6, are not significantly different, the hybrid approach guarantees the physical trend of the MFB50, which the ANN model is not always able to follow. Furthermore, the size of the dataset used to develop

the hybrid approach (704 points) was smaller than that used for the purely ANN-based approach (1200 points), thereby reducing the time and effort required for the experimental tests. Regarding the trade-off between the accuracy/ability to capture the physical trend of the phenomenon and the size of the data required, the hybrid approach has an advantage with respect to the purely ANN-based one and therefore is preferred in this work. Only for the case of the homologation cycle does the ANN have slightly better accuracy, and this is because the engine points included during the cycle are within the domain of the data used to train the ANN. It is well known that, for a given dataset, ANNs have higher accuracy than other approaches. At the same time, the advantage of the hybrid approach can be further highlighted through the results of Test 1, which include data with changes in many variables, and this approach is able to better capture these effects.

Table 6. Results for MFB50.

Type of Test	Pure ANN RMSE [°CA]	Hybrid Approach RMSE [°CA]
Spark Sweep	1.74	1.43
λ Sweep	2.30	1.72
VVT Sweep	3.19	2.38
Tans Sweep	0.92	0.50
Test 1	9.91	7.07
Homologation Cycle	4.51	5.27
Track Cycle	2.46	2.47

Figure 23 shows some results achieved for steady-state and transient conditions. The upper part of the figure refers to the steady-state result, which shows that the hybrid approach has slightly higher accuracy than the pure ANN model. Meanwhile, the lower part of the graph refers to the results obtained for the MFB50 during a homologation cycle, as represented in Figure 6. The lower-left plot in Figure 21 refers to the results during the complete cycle, whereas the lower-right plot is a magnified version of this complete cycle. As can be seen, the yellow line (referring to the hybrid approach) is almost always able to closely follow the experimental signal (blue line), which is not the case with the pure ANN model (red line).

5.2. Pmax, IMEP, and Knock

The Pmax and IMEP models take one of the inputs as the simulated MFB50, which captures the effects of different actuations/variables, and the engine load. Meanwhile, the MAPO takes as input the simulated Pmax and MFB50, along with other inputs, such as the engine speed, load, and λ. Table 7 shows the results achieved for various steady-state and dynamic/transient tests. As seen in this table, the results achieved show accuracy with an acceptable level of error, with the RMSE being within 5% with respect to the maximum value of each index.

Table 7. Results for Pmax and IMEP.

Type of Test	Pmax RMSE [Bar]	IMEP RMSE [Bar]	MAPO RMSE [Bar]
Spark Sweep	3.03	0.84	0.89
λ Sweep	3.11	0.73	0.96
VVT Sweep	5.31	1.19	0.71
Tans Sweep	2.73	0.55	1.70
Homologation Cycle	2.02	0.24	0.08
Track Cycle	4.55	1.14	1.10

Figure 24 shows a visual representation of some results obtained under the steady-state tests. For the steady-state tests, the knock model evaluated the MAPO98.5 following the model explained in the Knock Model section (Section 4.4).

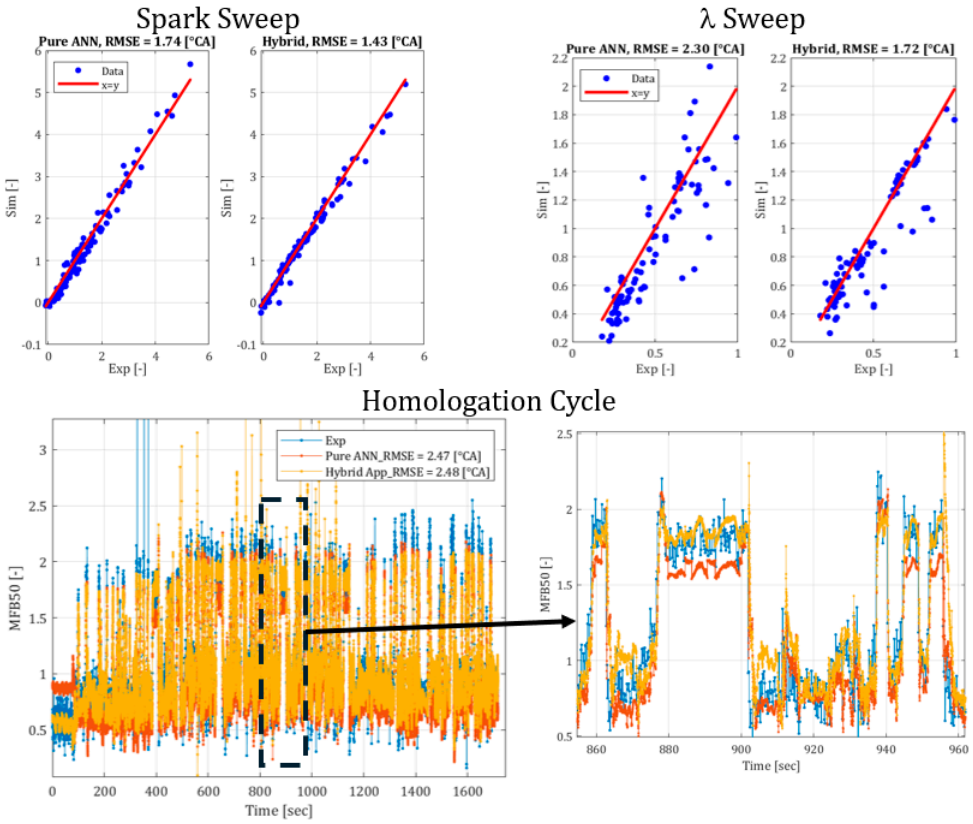


Figure 23. Results for MFB50 both under steady-state and transient conditions.

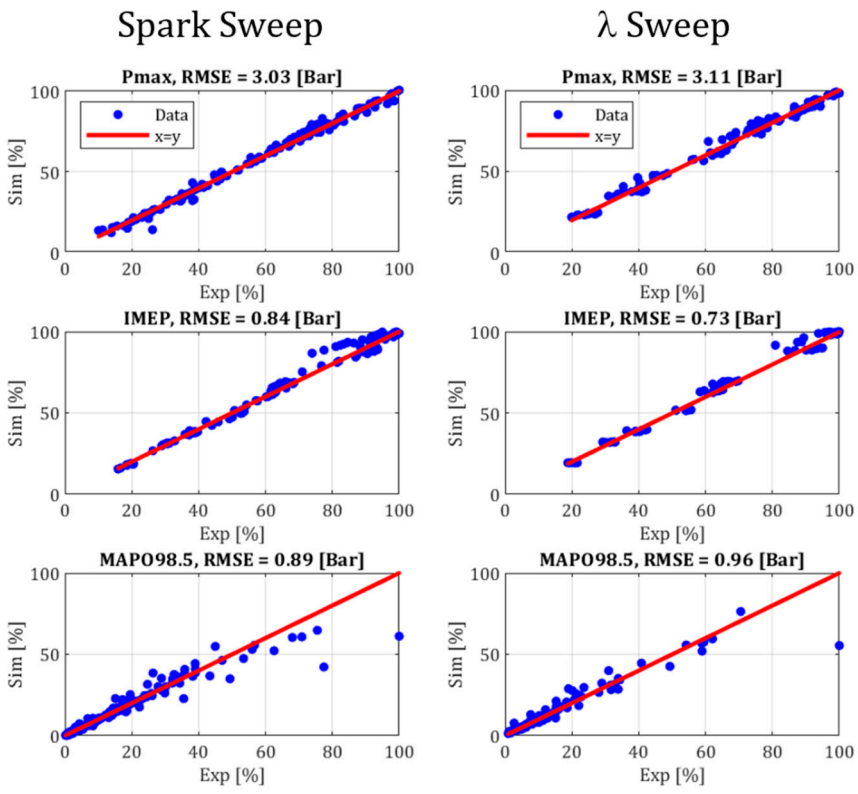


Figure 24. Pmax, IMEP, and MAPO results under steady-state tests.

Meanwhile, Figure 25 shows the results achieved for a homologation cycle; the left graph shows the complete cycle, while the right graph shows the magnified version. For

these dynamic/transient tests, the knock model evaluated the cyclic MAPO by extracting the logarithmic random values using the μ and σ simulated by the ANN model, which reflected the same probability distribution as the measured values of the MAPO.

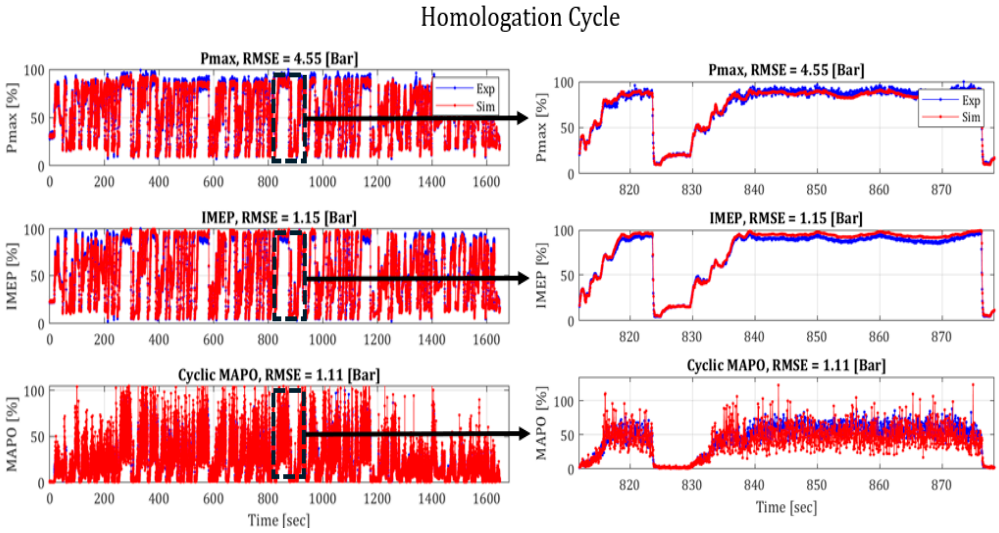


Figure 25. Pmax, IMEP, and MAPO results under transient tests.

5.3. Exhaust Gas Temperature

The exhaust gas temperature model does not consider the effect of the thermocouple behavior under steady-state tests. The thermocouple model intervenes only in the case of transient tests. Table 8 shows the results achieved for this index for both the ANN-based and hybrid approaches under both steady-state and dynamic tests. As seen in the table, the results achieved when using the hybrid approach are similar to and/or better than those for the ANN-based approach and guarantee the capture of the physical trend of the phenomenon.

Table 8. Results for exhaust gas temperature.

Type of Test	Pure ANN RMSE [°C]	Hybrid Approach RMSE [°C]
Spark Sweep	11.40	19.18
λ Sweep	10.47	14.19
VVT Sweep	8.55	30.74
Tans Sweep	9.96	12.47
Test 1	74.36	62.46
Homologation Cycle	55.66	32.11
Track Cycle	15.81	11.62

The upper plots in Figure 26 show the results achieved for the exhaust gas temperature during a steady-state test. During these tests, the thermocouple model did not intervene, and the results shown are the outputs of the models (the hybrid approach, Figure 22, and the purely ANN-based model). Meanwhile, the lower plots in Figure 26 refer to the simulations carried out on a homologation cycle (represented in Figure 6), where the output of the exhaust gas temperature models became the input to the thermocouple model (as described in Section 4.5.4), and the results shown are the outputs from the thermocouple model. The lower-left plot in Figure 26 presents the results for the entire cycle, while the lower-right plot provides a magnified view of this full cycle. As seen in this figure, the hybrid approach (yellow line) follows more closely the measured value of the temperature (blue line), thereby further highlighting the advantages of following this approach.

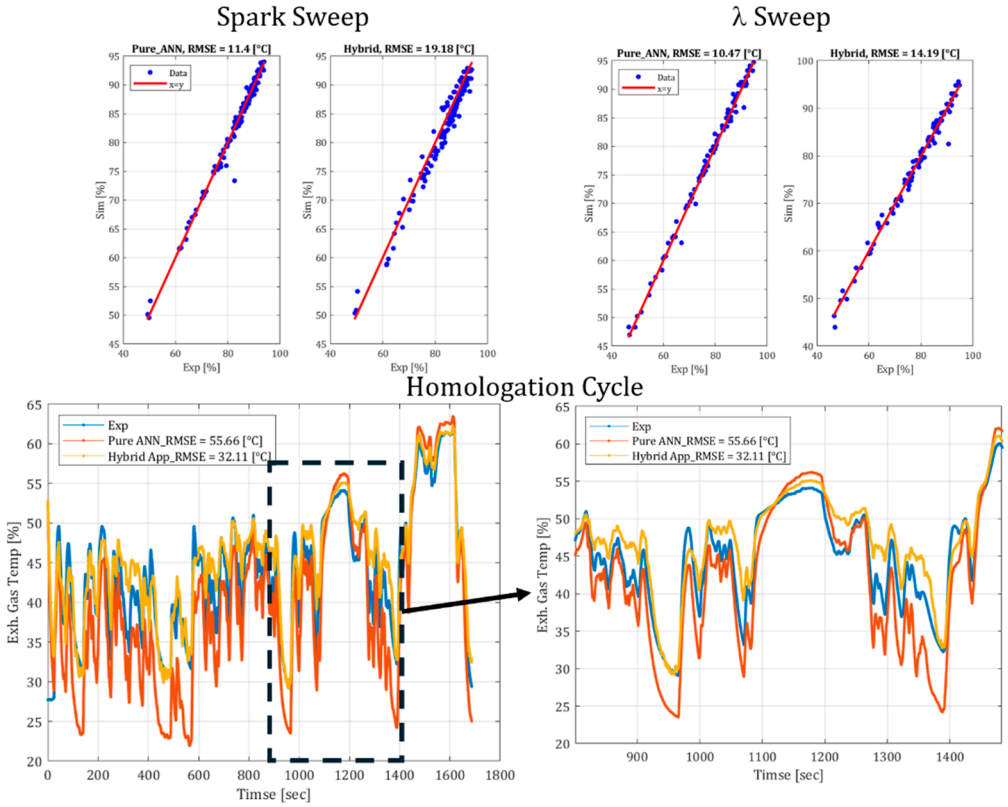


Figure 26. Exhaust gas temperature results under transient tests.

6. Conclusions and Future Work

This study highlights the significant advancements achieved through the integration of traditional analytical methods with modern machine learning techniques in engine modeling. The hybrid approach developed in this work is based on the effects superposition principle, and it combines the precision and robustness of physical models with the adaptability and learning capabilities of artificial neural networks. By incorporating analytical corrective functions into ML-based engine models, this study addresses key limitations of the current modeling approaches, such as feature extrapolation, parameter estimation, and the ability to capture physical dynamics. The main innovative contribution of this work is the extension of the hybrid approach to model complex physical problems, even with a larger number of independent variables. In particular, artificial neural networks are used to capture the effects of variables for which the operating range is well defined (such as the engine speed and load), while analytical functions are applied to describe and extend the trends of the output variables (such as the combustion, knock, and exhaust gas temperature indices) with respect to those independent variables that could assume values that differ from the calibrated ones. This allows us to increase the extrapolation capabilities of a standard data-driven model, and it does not have an impact on the effort required for the collection of the experimental data.

This experimental campaign and the development of a complete engine simulator demonstrate the superior predictive accuracy and generalization performance of hybrid models compared to purely data-driven methods. These enhanced models maintain consistency with physical trends, even under operating conditions that deviate significantly from the training data. This capability is crucial in ensuring the reliability and applicability of engine simulators in real-world scenarios, including under anomalous engine operations and failure predictions. It is demonstrated that such an approach can be effectively applied to combustion and knock modeling, but also to exhaust gas temperature estimation, through the implementation of a semi-physical thermocouple model. Moreover, an additional benefit of the proposed approach is the possibility to define the values of the calibratable

coefficients with fewer engine points, with respect to a purely data-driven approach. This results in a reduction in the computational effort required for the calibration of the model and in the possibility to extend the data collected at the test bench via simulations, with a concrete reduction in costs.

On the other hand, tests for the calibration of the analytical functions have to be carried out by varying a single variable and keeping all other variables constant, so as to isolate the effect of each input feature on the calculated output. This, in some cases, leads to variation in the standard procedures to conduct the campaign to calibrate the main ECU control strategies. At the same time, as mentioned above, the number of engine points required for complete model calibration is smaller than that needed for a purely data-driven methodology.

Future research directions include further improving the accuracy and robustness of the data-driven models and exploring additional applications of the hybrid approach to model different engine parameters, such as functional indices. Moreover, the developed engine simulator will be applied for offline engine calibration activities and to design and pre-evaluate new durability tests and driving cycles.

Author Contributions: Conceptualization, supervision, validation, writing-review & editing, G.G. and I.K.; Writing—original draft, conceptualization, data curation, visualization, A.B., F.P.S. and B.P.; Supervision, N.C. and D.M. All authors have read and agreed to the published version of the manuscript.

Funding: Financed by the European Union—NextGenerationEU through the Italian Ministry of University and Research under PNRR—Mission 4 Component 2, Investment 3.3 "Partnership extended to universities, research centers, companies and founding of basic research projects" D.M. 352/2022—CUP J33C22001350009.



Data Availability Statement: Authors cannot share the data used for this work.

Conflicts of Interest: Authors Giovanni Galasso and Ioannis Kitsopanidis were employed by the company Ferrari S.p.A. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
ANN	Artificial Neural Network
ECU	Engine Control Unit
IMEP	Indicated Mean Effective Pressure
λ	Lambda
LUT	Look-Up Table
MAPO	Maximum Amplitude Pressure Oscillation
MBD	Model-Based Learning
MFB50	50% of Mass Fraction Burnt
ML	Machine Learning
PDF	Probability Density Function
Pmax	Maximum In-Cylinder Pressure
R2	R-Square
RMSE	Root Mean Square Error
RON	Research Octane Number
SA	Spark Advance
Tans	Intake Manifold Temperature
VVT	Variable Valve Timing

References

1. Monteiro, N.B.R.; da Silva, E.A.; Neto, J.M.M. Sustainable development goals in mining. *J. Clean. Prod.* **2019**, *228*, 509–520. [CrossRef]
2. AVL Public Discussion. Available online: https://www.avl.com/documents/4329920/48266926/AVL+Emission+Test+System+and+Emission+New+Regislation.pdf (accessed on 17 November 2021).
3. DieselNet. Emission Standards: Summary of Worldwide Engine and Vehicle Emission Standards, DieselNet. Available online: https://dieselnet.com/standards/ (accessed on 11 May 2021).
4. Shamekhi, A.M. Engine Model-Based Pre-calibration and Optimization for Mid-level Hierarchical Control Design. *SAE Int. J. Engines* **2021**, *14*, 651–669. [CrossRef]
5. Şener, R.; Gül, M.Z. Optimization of the combustion chamber geometry and injection parameters on a light-duty diesel engine for emission minimization using multi-objective genetic algorithm. *Fuel* **2021**, *304*, 121379. [CrossRef]
6. Giovannardi, E.; Brusa, A.; Petrone, B.; Cavina, N.; Corti, E.; Barichello, M. An Enhanced Light Gradient Boosting Regressor for Virtual Sensing of CO, HC and NOx. In Proceedings of the 2023 IEEE International Workshop on Metrology for Automotive (MetroAutomotive), Modena, Italy, 28–30 June 2023; pp. 1–6. [CrossRef]
7. Brusa, A.; Giovannardi, E.; Barichello, M.; Cavina, N. Comparative Evaluation of Data-Driven Approaches to Develop an Engine Surrogate Model for NOx Engine-Out Emissions under Steady-State and Transient Conditions. *Energies* **2022**, *15*, 8088. [CrossRef]
8. Petrone, B.; Giovannardi, E.; Brusa, A.; Cavina, N.; Kitsopanidis, I. *Development of an Automatic Pipeline for Data Analysis and Pre-Processing for Data Driven-Based Engine Emission Modeling in a Real Industrial Application*; SAE Technical Paper 2024-01-2018; SAE International: Warrendale, PA, USA, 2024. [CrossRef]
9. Shethia, F.P.; Mecagni, J.; Brusa, A.; Cavina, N. *Development and Software-in-the-Loop Validation of an Artificial Neural Network-Based Engine Simulator*; SAE Technical Paper 2022-24-0029; SAE International: Warrendale, PA, USA, 2022. [CrossRef]
10. Brusa, A.; Mecagni, J.; Corti, E.; Silvestri, N. Application of a Neural-Network-Based Algorithm for the Real-Time Correction of the In-Cylinder Pressure Signal Sensed with a Piezoelectric Washer. *SAE Int. J. Engines* **2023**, *16*, 663–679. [CrossRef]
11. Fossier, S.; Robic, P.-O. Maintenance of complex systems—From preventive to predictive. In Proceedings of the 2017 12th International Conference on Live Maintenance (ICOLIM), Strasbourg, France, 26–28 April 2017; pp. 1–6. [CrossRef]
12. Brusa, A.; Mecagni, J.; Cavina, N.; Corti, E.; Cucchi, M.; Silvestri, N. *Development and Experimental Validation of a Control-Oriented Empirical Exhaust Gas Temperature Model*; SAE Technical Paper 2020-24-0008; SAE International: Warrendale, PA, USA, 2020. [CrossRef]
13. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
14. Martin, D.R.; Rocci, B. *Virtual Exhaust Gas Temperature Measurement*; SAE Technical Paper 2017-01-1065; SAE International: Warrendale, PA, USA, 2017. [CrossRef]
15. Finesso, R.; Spessa, E.; Yang, Y.; Conte, G.; Merlino, G. *Neural-Network Based Approach for Real-Time Control of BMEP and MFB50 in a Euro 6 Diesel Engine*; SAE Technical Paper 2017-24-0068; SAE International: Warrendale, PA, USA, 2017. [CrossRef]
16. Brusca, S.; Lanzafame, R.; Messina, M. *A Combustion Model for ICE by Means of Neural Network*; SAE Technical Paper 2005-01-2110; SAE International: Warrendale, PA, USA, 2005.
17. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
18. Warey, A.; Gao, J.; Grover, R. Prediction of Engine-Out Emissions Using Deep Convolutional Neural Networks. *SAE Int. J. Adv. Curr. Pr. Mobil.* **2021**, *3*, 2863–2871. [CrossRef]
19. Lee, S.-Y.; Andert, J.; Pischinger, S.; Ehrly, M.; Schaub, J.; Koetter, M.; Ayhan, A.S. *Scalable Mean Value Modeling for Real-Time Engine Simulations with Improved Consistency and Adaptability*; SAE Technical Paper 2019-01-0195; SAE International: Warrendale, PA, USA, 2019. [CrossRef]
20. Brusa, A.; Cavina, N.; Rojo, N.; Mecagni, J.; Corti, E.; Ravaglioli, V.; Cucchi, M.; Silvestri, N. Development and Experimental Validation of an Adaptive, Piston-Damage Based Combustion Control System for SI Engines: Part 1—Evaluating Open-Loop Chain Performance. *Energies* **2021**, *14*, 5367. [CrossRef]
21. Brusa, A.; Cavina, N.; Rojo, N.; Mecagni, J.; Corti, E.; Moro, D.; Cucchi, M.; Silvestri, N. Development and Experimental Validation of an Adaptive, Piston-Damage-Based Combustion Control System for SI Engines: Part 2—Implementation of Adaptive Strategies. *Energies* **2021**, *14*, 5342. [CrossRef]
22. Brusa, A.; Shethia, F.P.; Mecagni, J.; Cavina, N. Advanced, Guided Procedure for the Calibration and Generalization of Neural Network-Based Models of Combustion and Knock Indexes. *SAE Int. J. Engines* **2024**, *17*, 153–164. [CrossRef]
23. Cavina, N.; Brusa, A.; Rojo, N.; Corti, E. *Statistical Analysis of Knock Intensity Probability Distribution and Development of 0-D Predictive Knock Model for a SI TC Engine*; SAE Technical Paper 2018-01-0858; SAE International: Warrendale, PA, USA, 2018. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.