Systematic assessment of data-driven approaches for wall heat transfer modelling for LES in IC engines using DNS data

Author(s):
Impagnatiello, Matteo; Bolla, Michele; Keskinen, Karri; Giannakopoulos, George; Frouzakis, Christos E.; Wright, Yuri; Boulouchos, Konstantinos

Publication date:
2022-02

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-b-000514668

Originally published in:
International Journal of Heat and Mass Transfer 183(Part B), https://doi.org/10.1016/j.ijheatmasstransfer.2021.122109

International Journal of Heat and Mass Transfer 183 (2022) 122109

Contents lists available at ScienceDirect

International Journal of Heat and Mass Transfer

journal homepage: www.elsevier.com/locate/hmt

Systematic assessment of data-driven approaches for wall heat transfer modelling for LES in IC engines using DNS data

Matteo Impagnatiello*, Michele Bolla, Karri Keskinen, George Giannakopoulos, Christos E. Frouzakis, Yuri M. Wright, Konstantinos Boulouchos

Swiss Federal Institute of Technology (ETH Zurich), Laboratory for Aerothermochemistry and Combustion Systems, Sonneggstrasse 3, CH-8092 Zurich, Switzerland

ARTICLE INFO

Article history:
Received 27 May 2020
Revised 15 March 2021
Accepted 13 October 2021

Keywords:
Data-driven modelling
Wall heat transfer
Wall modelling
LES
DNS
IC engines

ABSTRACT

Data-driven (DD) methods offer a promising pathway towards novel modelling solutions in fluid flow and heat transfer. In this study, we investigate the application of DD neural network (NN) methods on wall heat transfer modelling in the context of wall-modelled large-eddy simulation (WMLES) in engines, focusing on the systematic evaluation of criteria for the successful DD model generation. High-fidelity input data for model training and testing is generated by spatial filtering of DNS and wall-resolved LES fields in several engine and engine-like configurations. The NN-based models are constructed using different input data and wall-adjacent cell schemes, while cell size and network complexity are also varied. The evaluated NN-based models demonstrate improved performance with respect to classical wall functions, indicating promising potential for engineering applications. In particular, better modelling results were obtained with the inclusions of a wall-normal cell Reynolds number and of data from the second wall-normal cell. Such a two-cell input format appears to offer a good compromise between performance and complexity. Both the present NN models and literature reference approaches generally perform better in unburned regions than in burned ones. In near-wall regions with flame fronts, we present an analysis dividing samples into “unburned”, “burned”, and “flame boundary” zones exposing different characteristics and a varying degree of modelling difficulty.

1. Introduction

In recent years, data-driven methods based on machine learning and deep learning (DL) have attracted a lot of interest in fluid dynamics problems [1,2]. Data-driven models are particularly attractive for describing processes where there are no obvious model closure equations. For computational fluid dynamics (CFD), in addition to the challenging identification of the model closures that are well-suited for a statistical description, the application of DD methods hinges upon the availability of high-fidelity data which are typically difficult to obtain with sufficient temporal and spatial resolution. For Large-Eddy Simulation (LES) the requirements in terms of high-fidelity data needs are higher than for Reynolds–Averaged Navier–Stokes (RANS) because data is required at spatially and temporally-resolved scales. Neural networks (NN) have been successfully applied for subgrid scale (SGS) modelling for LES using Direct Numerical Simulation (DNS) data [3–6]. Overall, a trend of employing increasingly advanced network architectures is observed, for example with physics-informed networks “perceiving” physics-based properties such as conservation of mass and momentum [7].

Wall-modelled LES (WMLES) is a modern, highly relevant approach for engineering applications wherein the considerable near-wall resolution requirements of LES are dramatically reduced, hence extending the Reynolds number range of scale-resolving simulations [8]. The wall models attempt to predict wall shear stress and wall heat flux, and are often based on simplified versions of the governing equations or equivalent algebraic profiles (wall functions). Although several novel approaches have been recently proposed (see for example [9,10]), such developments are often focused on wall shear stress rather than on heat transfer. While equilibrium assumptions may be applicable in certain cases, many complex flows deviate considerably therefrom. In internal

* Corresponding author.
E-mail address: matteoi@ethz.ch (M. Impagnatiello).

https://doi.org/10.1016/j.ijheatmasstransfer.2021.122109

combustion (IC) engine flows, deviation from classical near-wall flow profiles has been observed both experimentally [11,12] and numerically [13]. In such flows, wall models incorporating non-equilibrium effects [14,15] have demonstrated promising improvements to classical models based on wall functions. However, despite known deficiencies, classical wall functions (e.g. [16,17]) or even no-model approaches (e.g. [18,19]) are frequently employed in LES of IC engines.

Assessing the correctness of IC engine wall models locally is non-trivial due to limited availability of high-fidelity data. Only few experimental IC engine near-wall velocity measurements exist [11,12], and the authors are not aware of any high-fidelity time-resolved descriptions of wall heat flux and concurrent near-wall temperature profiles. The recent emergence of DNS and wall-resolved LES data in real IC engines [20,21] and engine-like configurations [13] enables the detailed investigation of boundary layers; the latter data have been successfully employed to assess wall modelled LES approaches [22]. Even in this case, however, systematic statistical analyses are challenged by (i) the scarcity of spatially homogeneous regions in the combustion chamber, (ii) the considerable temporal variation within the engine cycles, and especially (iii) the prevalence of chemical reactions in situations with the highest heat fluxes. While the construction of accurate wall models for IC engines is challenging, it is also considered critical by both academic and industrial communities [22,23].

In the context of channel flow LES, NN applications with promising perspectives start to appear for momentum boundary layers [24,25] and wall heat flux [26]. A first data-driven step for IC engines was recently undertaken by the authors with the introduction of an algebraic wall heat transfer (WHT) model [27].

1.1. Study objectives

In the current work, we present a data-driven neural network WHT model for WMLES, developed using high-fidelity simulation data from both real IC engines (operated under motored as well as fired conditions) and engine-like configurations. To the best of our knowledge, this is the first attempt of using a neural network for the development of local wall heat flux models for LES of IC engines. In addition, the main contribution of the work is to examine the factors impacting the functionality of such models by addressing the considerations that should be made when applying data-driven approaches to develop models for complex flows such as those encountered in IC engines.

We present a systematic assessment of various aspects related to the chosen data-driven approach and its successful development, addressing:

1. The influence of data selection and choice of input features (both flow variables and their location).
2. How to consistently compare performance of different neural network topologies.
3. How to assess performance for different system states (reactive and non-reactive).

The performance indicators of such seemingly abstract parameters can be connected with the underlying flow physics, such as the spatial distance of information exchange (i.e. the integral length scale). Beyond IC engines, the present analyses and parameter sensitivities are useful for the description of complex flow and heat transfer processes with data-driven methods.

The paper is structured as follows: first, the data sets for model generation and benchmarking are presented in Section 2. Special attention is paid to a meaningful division into training and testing sets, in order to appropriately account for different fluid flow and thermochemical conditions. Section 3 describes the basics of WHT modelling and presents existing methods that are later used as a reference. Section 4 describes the construction of artificial neural network WHT models, including input feature configuration and network choice. Section 4.4 reports how the generated networks perform in WHT modelling and discusses how the performance depends on network complexity, thermochemical (unburned/burned) state, and WMLES filter specifications (grid size and aspect ratio). Finally, Section 5 provides answers to the research questions.

2. Data and pre-processing

As the choice of data plays a pivotal role in DD model development and only few wall-resolved simulations in real IC engines or engine-like, confined geometries exist at present, this section summarizes the respective configurations and datasets, and provides a short description of the numerical platforms used to obtain them. For details, the reader is referred to the corresponding references. Filtering procedures applied to the datasets are subsequently detailed, in addition to the data splitting into training and testing sets. Finally, to reduce the sensitivity of performance evaluation metrics, clustering of the samples by their thermochemical state is discussed.

2.1. Input data

In data-driven models, availability and quality of the input data have a primary influence on model performance. Recent DNS and wall-resolved LES simulation results from three setups are used in this study, which were generated with the spectral element code Nek5000 [28]. The setups consist of one canonical turbulent premixed flame in a confined cylindrical geometry and two IC engines: the TCC-III optical engine [29] and the Darmstadt optical engine [12] (Fig. 1). In the rest of the paper, the datasets will be referred to as TCC, Syngas, TUD-NR and TUD-R with the last two sharing the same engine geometry under non-reactive (i.e. motorized operation) and reactive (i.e. fired operation) conditions. A brief description of the datasets follows.

2.1.1. TCC dataset

The TCC dataset is obtained from multi-cycle high-fidelity wall-resolved LES [21] of the TCC-III two-valve pancake engine which has a bore of $D = 92$ mm and stroke $H = 86$ mm. It is operating at motored conditions with intake pressure $P_{intake} = 0.95$ bar and at an engine speed 500 rpm; the walls are considered isothermal at a temperature $T_w = 500$ K. The valve timing of the simulated operating conditions has been slightly modified compared to the experimental one [29] to facilitate multiple-cycle simulations with the low Mach number solver Nek5000 by avoiding valve overlap [21]. The samples for model development are extracted from a 32×32 mm^2 square centrally-located on the piston surface and a 16×64 mm^2 rectangle on the cylinder head between the valves. Nine evenly spaced time-steps between -40° CA and TDC are considered. To improve the statistics, the data are collected from five engine cycles. As the datasets on the piston and head statistically showed only minor differences in terms of filtered wall heat transfer PDF (particularly close to TDC) — probably due to the pancake shape — both datasets have been merged into a single dataset that will be denoted as TCC.

2.1.2. Syngas dataset

Syngas data at reactive conditions are obtained from the DNS of a lean ($\phi = 0.3$) syngas-air turbulent premixed flame (mole fraction ratio $X_{CO} : X_{H2} = 3$) in a closed cylindrical domain ($D =$ 75 mm, $H = 30$ mm) designed to emulate an engine geometry during the late compressions stroke [30]. The initial temperature and pressure are 820 K and 1 bar and the no-slip inert walls are isothermal ($T_w = 550$ K). Homogeneous isotropic turbulence with

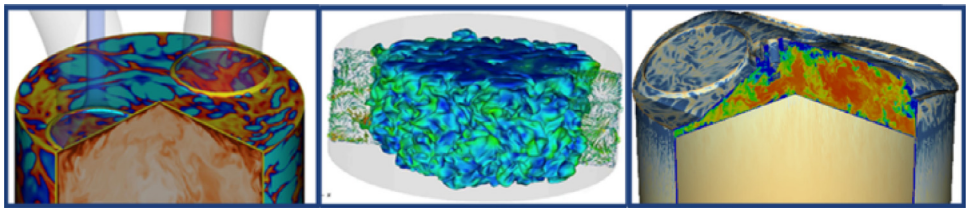

Fig. 1. The three configurations used to generate the data. From left to right: TCC III, Syngas, and TU Darmstadt. The images are mainly illustrative of the configuration and the quantities shown are: TCC III: wall heat flux and temperature; Syngas: Flame iso-surface coloured by velocity; TU Darmstadt: wall heat flux and temperature.

two initial turbulence intensities ($u'/S_L = 2$ and 6 and $l/\delta_f = 2$) are considered [31] where $\delta_f = 1.1$ mm and $S_L = 1.38$ m/s are the laminar flame thickness and speed, respectively. The mixture is ignited using the equilibrium solution in a centrally-located sphere with a diameter $d = 4.5$ mm. The sampling surface is a 50×50 mm^2 central square on the top and bottom walls. The data are extracted at different times as the flame front proceeds towards the walls and after the head-on wall impingement so that samples both in the presence and absence of the flame in the wall vicinity are considered in the Syngas dataset.

2.1.3. TUD-NR and TUD-R datasets

The TU Darmstadt (TUD) data are obtained from the DNS of a single compression stroke in the TUD four-valve pent-roof optical engine ($D = 86$ mm, $H = 86$ mm) under both motored and fired conditions with $P_{intake} = 0.4$ bar operating at 800 rpm. The conditions correspond to test case B of the experimental test matrix reported in Ding et al. [12]. Both motored and fired DNS runs have been initialized at inlet valve closure (IVC) (-130° CA), where the initial fields of velocity, temperature and composition have been obtained from a dedicated auxiliary multi-cycle LES calibrated against experimental data. For a detailed description of the DNS-LES simulation workflow the reader is referred to [20]. Under motored conditions the working fluid is pure air and the in-cylinder wall temperature is fixed to 333 K. Under fired conditions, the intake composition is a homogeneous stoichiometric CH4-air mixture and the measured wall temperature has been imposed. For the piston surface a constant temperature of 570 K has been prescribed. The spark timing has been set at -22.2° CA and the mixture is ignited using the equilibrium solution in a sphere located between the spark-plug electrodes with a diameter $d = 1.5$ mm. In order to keep the computational cost at an affordable level (around 10 million core-hours for the advancement of the reactive simulation from -22.2° CA to *TDC*), a two-step reduced chemical mechanism [32] has been employed, with the pre-exponential factors and activation temperatures adjusted to replicate the laminar flame speed and thickness as well as the adiabatic flame temperature that are computed using the detailed GRI3.0 mechanism [33] over the range of thermochemical conditions encountered during compression.

The samples are extracted from a centrally-located 36×36 mm^2 square on the piston surface. Under non-reactive conditions (TUD-NR dataset) nine evenly-spaced timesteps between -40° CA and *TDC* are considered, while under fired conditions (TUD-R) the nine timesteps are clustered between spark time (-22.2° CA) and *TDC*.

2.2. Data filtering

For all configurations, the near-wall DNS solution from the spectral element code Nek5000 is mapped onto a uniform grid at DNS resolution by means of a high-order spectral interpolation. Subsequently, a density-weighted top-hat filtering operation is performed to produce data at LES resolution, mimicking the loss of information at the sub-grid scale. This approach, also called "a-priori" approach, is common practice in literature. The emulated LES grid consists of spatially uniform hexahedral control volumes where the wall-normal cell size is half the wall-parallel directions. This leads to an aspect ratio (defined as the wall-normal to wall-parallel ratio) of 1:2 that mimics a one-level wall-normal refinement to be expected in practice. All models are trained using only the reference filter aspect ratio of 1:2. The performance of the models is subsequently assessed with the testing datasets using different filter aspect ratios (1:1, 1:2, 1:4, 1:6, and 1:8) in order to assess model robustness. Throughout this paper the filter size refers to the wall-normal direction. Since LES are inherently dependent on filter size, and the contribution of the sub-grid part increases with it, all datasets are filtered with 21 equally-spaced filter sizes in order to accommodate the range of typical engineering-type LES simulations. In particular, filter sizes between 0.128 mm and 0.768 mm are employed for the TCC, TUD-NR, and TUD-R datasets while filter sizes between 0.4 mm and 2.4 mm are considered for the Syngas dataset.

2.3. Data splitting: training and test sets

The samples are split into a training set, used only to tune the model parameters, and a test set, used only for performance evaluation. The availability of data from different setups allows for a dataset-specific splitting. We choose to include both non-reactive and reactive data in the training phase, as their thermo- and fluid-dynamic quantities are highly contrasting.

To support the split strategy, we first consider the statistical distribution of the datasets in terms of representative thermodynamic conditions used as explaining variables (or descriptive quantities). The main purpose of this preliminary screening is to assess how widespread the conditions within and between the datasets are.

Here, we select the filtered density and temperature in the first wall-adjacent cell as descriptive quantities as illustrated in Fig. 2 by means of mono-chromatic, clipped joint-pdf (JPDF). The curves from low to high density for each dataset represent separate time instances. Along a single curve, the increase in filtered temperature is the result of enlarged filter size.

Overall, the spread of the samples in filtered $T - \rho$ space across a wide range of temperatures and densities (pressures) is beneficial for assessing the generalization for both model training and testing. Additionally, these variables can be considered in combination with the probability distribution of the target variable (i.e. the wall heat flux) as shown in Fig. 3. The clustering of samples as non-reactive, "flame boundary" and "burned" zone will be introduced in Section 2.4.

Under non-reactive conditions, the filtered heat flux distributions show a positively skewed PDF with peak values (as represented by the 99th percentile of the data) of around 400 kW/m^2 and 200 kW/m^2 for the TCC and TUD setups, respectively. For the Syngas dataset, a bi-modal distribution is observed with a clear distinction between situations where the turbulent flame is far

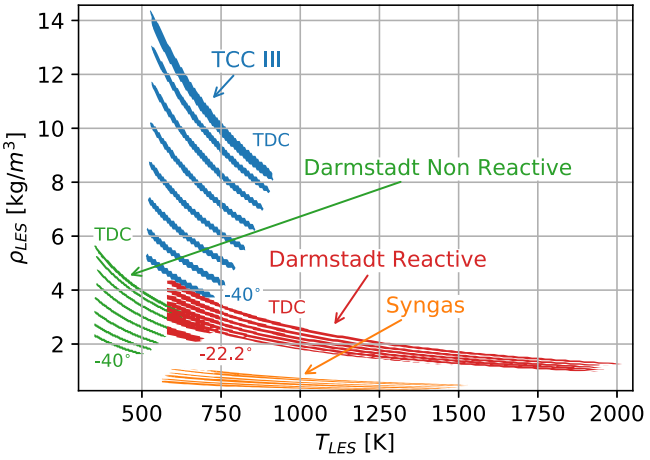

Fig. 2. Sample distribution of filtered temperature and density of first LES wall-adjacent cell coloured by dataset. Each line represents a different time instance.

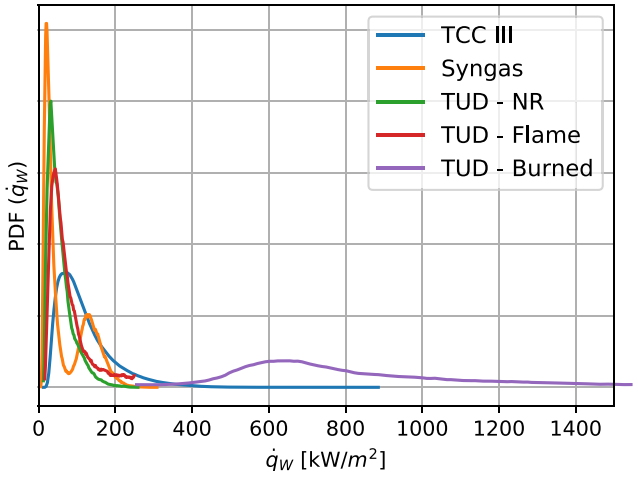

Fig. 3. Distribution of filtered wall heat flux for different datasets.

Table 1
Data splitting across training and testing.

	NON-REACTIVE	REACTIVE
TRAINING	TCC	Syngas
TEST	TUD-NR	TUD-R

from or in the vicinity of the wall (i.e. within the first LES cell). The peak value of around 250 kW/m^2 under reactive conditions is comparable to the TCC and TUD setups under motored conditions. The main reason for the rather low heat transfer in the Syngas case is the low initial pressure of 1 bar (the peak pressure is approximately 2 bar) in combination with the low equivalence ratio ($\phi = 0.3$), resulting in a low adiabatic flame temperature (close to 1850 K).

The TUD-R (Burned) dataset exhibits significantly higher filtered wall heat flux values (approximately 2 MW/m^2) compared to the other datasets. There is a factor of 5–10 between non-reactive and reactive peak heat flux values.

From the aforementioned considerations, a splitting is proposed according to Table 1, in order to introduce two challenging features during the testing phase. First, contrary to the TCC case, the pent-roof geometry of the TUD engine induces significant tumble flow and although the ranges of WHT values in the TUD-NR and TCC datasets are comparable, the different flow pattern has a significant impact on the boundary layer characteristics. In addition, under fired operation (TUD-R dataset) the wall heat flux is significantly higher compared to the other datasets.

2.4. Clustering between burned, unburned and flame boundary regions

To improve the overall performance evaluation procedure, samples from the testing datasets are grouped into three clusters denoted as: “unburned”, “burned”, and “flame boundary” according to their thermochemical state. This grouping allows for a more targeted assessment of the specific situations that are different in nature, as can be depicted from the profound differences in local heat fluxes. Furthermore, sensitivity of the performance evaluation metric to the share of (“burned” or “unburned”) samples can also be reduced.

All samples from the TUD-NR dataset can be clearly assigned to the “unburned” group. For the TUD-R, the turbulent flame impingement on the engine piston and walls is a highly transient process where the flame approaches the boundaries in different directions [34]. A clear distinction between “unburned” and “burned” situations is not obvious in engine LES since at high pressures the laminar flame thickness is typically smaller than the LES cell size and a turbulent combustion model is required to estimate the subgrid thermochemical state and chemical reactions.

In this work, we do not concentrate on the specifics of the coupling with the combustion model. Instead, we conceptualize the flame-wall interaction (FWI) phenomenology into a simplistic state-based classification as “unburned”, “flame boundary”, and “burned”. The distinction between the zones is based on the filtered progress variable, which is computed using the local filtered oxygen mass fraction. Values of 0 and 1 correspond to “unburned” and “burned” states, respectively, while intermediate values denote the “flame boundary”. A tolerance of 0.01 was applied for the limits of the “flame boundary” (i.e. 0.01–0.99).

The “unburned” and “burned” regions have in common that there are no sub-grid fluctuations of the progress variable. This means that for the two extrema the same model can be formulated without an explicit tracing of the progress variable.

For the intermediate values of progress variable (“flame boundary”) the situation is different. As the flame approaches the wall, its impact is detected in the filtered near-wall cell (as increased temperature) before the actual high-temperature front reaches the wall. Hence, using only the quantities listed above may inherently lead to heat flux overprediction. This advanced flame sensing makes the “flame boundary” region the most difficult to model. Note that the advanced flame sensing is mainly expected to appear in head-on quenching circumstances.

The development of a more advanced model is not the focus of this work, and we limit ourselves to assessing the errors of the different models in the “flame boundary”. This can be justified by the overall small contribution of this intermediate region, accounting for less than 2% of the total wall heat flux for the test case considered here.

The clusters extracted from the TUD-R dataset are exemplarily depicted in Fig. 4, showing wall heat flux for DNS and for filtered DNS, and the spatial selection of “flame boundary” and “burned” samples. The “unburned” samples from the TUD-R dataset are deliberately removed to avoid the risk of including samples with spurious effects arising from remote flame sensing (this effect depends on the filter size). Preliminary tests indicated that the inclusion of a small amount of such biased samples has a significant effect on model performance. Therefore, “unburned” samples are exclusively taken from the TUD-NR dataset.

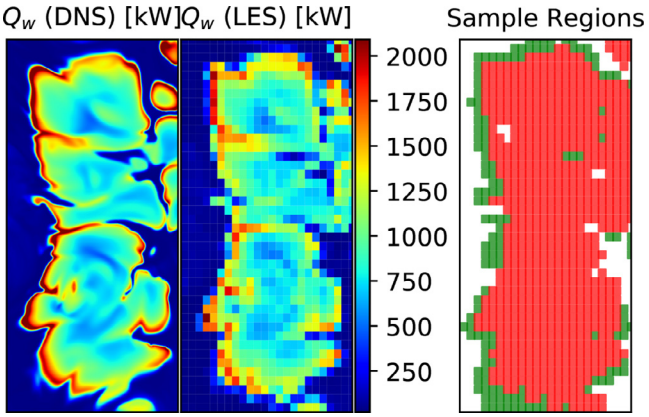

Fig. 4. Example of spatial distribution of wall heat flux for the TUD-Reactive case at −6° CA: DNS (left), LES (i.e. filtered DNS) with $\Delta_{LES} = 0.384$ mm (centre). Right is distribution of the corresponding sample classification (described in Section 2.4) into *Flame boundary* (green) and *Burned* (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Wall heat transfer modeling

3.1. Model formulation

Since the boundary layers in IC engines are not resolved in LES, the temperature gradient at the wall cannot be computed directly on the grid and a model is needed to reconstruct the effective gradient at the wall. We start from a generic expression for a WHT model linear with respect to the temperature gradient resolved by the LES grid:

$$Q_W = \lambda_W \cdot \left.\frac{dT}{dz}\right|_{W,LES} \cdot COEFF \tag{1}$$

where λ_W is the thermal conductivity at the wall temperature, T is the filtered temperature at the wall-adjacent node, and z is the wall-normal coordinate. $COEFF$ is a correction factor (i.e. the wall model) to account for the error made when evaluating $dT/dz|_W$ at the LES resolution. Correspondingly, for each grid tested, the effective COEFF values used for benchmarking are retrieved using DNS data as:

$$COEFF = \left(dT/dz|_{W,DNS}\right) \Big/ \left(dT/dz|_{W,LES}\right) \tag{2}$$

where the DNS near-wall gradient is surface-averaged. An important constitutive property of COEFF is that when the LES resolution approaches that of the DNS, COEFF will tend to unity making a wall model not necessary.

In this work, the proposed model for COEFF is purely data-driven and obtained using Artificial Neural Networks (ANN), which reconstruct the unknown function $COEFF = F(X)$, where X is a vector of inputs at the LES resolution as discussed in Section 4.

3.2. Reference models

Four additional approaches are considered as reference for the purpose of benchmarking the proposed data-driven models:

No model

The "no model" approach evaluates WHT using the resolved temperature gradient at the LES resolution, which is equivalent to setting $COEFF = 1$. This approach is employed as a reference since its results essentially quantify the need for a model. Inclusion of a reference/naïve model is usual in data-driven modelling to assist practitioners in judging model performance as the interpretability of these techniques is typically low.

WF: wall function

The second group of reference models is based on wall functions which express WHT in terms of resolved LES quantities avoiding the introduction of additional differential equations. The algebraic formulations may be based on empirical observations and/or simplifying assumptions of the boundary layer governing equations. In this work, the model proposed by Plengsaard and Rutland [35] is selected, because it was developed specifically for LES and IC engines. As in the case of various other models, heat flux and wall shear stress τ_W are linked in line with the Reynolds analogy. Its formulation is therefore explicitly dependent on the wall stress model since τ_W cannot be accurately computed at the LES resolution.

Other commonly used WF models [36,37] were also tested but are not discussed for brevity. Instead, only results for the WF model showing the best performance for the considered configurations are shown.

WF-τ_{DNS}: corrected wall Function

In contrast to the WF model where WHT is computed based on the wall shear stress, the WHT is computed here using τ_W extracted from the DNS. This variant of the model is used to discern the source of error arising from the inaccuracy of the τ_W model as opposed to the one of Q_W.

Algebraic model

The last reference approach is the algebraic model proposed in Bolla et al. [27]. This model has been calibrated using data from the aforementioned TCC and Syngas datasets, attempting to reconstruct COEFF and depending only on a single model parameter. A non-dimensional metric Δ^+ has been introduced in this model as a measure of how well the thermal boundary layer is resolved by the LES grid. The model is linear in Δ^+ and is formulated as:

$$COEFF = A \cdot \Delta^+ + 1 = 0.46 \cdot \frac{\rho_2}{\rho_w} \cdot \frac{T_1 - T_w}{T_2 - T_1} + 1 \tag{3}$$

where A is the constant trained on the data and the subscripts W, 1 and 2 refer to the wall, first and second wall-adjacent cells, respectively.

4. Artificial neural network

A Feed Forward Neural Network (multi-layer perceptron - MLP) is used in this work to map the set of input features (Section 4.2.1) to the single output (WHT). The Rectified Linear Unit (ReLU) is employed as activation function, and the loss function used during the training is the mean squared error (MSE) of the sample-based model-predicted heat flux. Gradient-based methods are usually employed when training this type of networks, the "Adam" [38] algorithm is used in this work.

4.1. Overview

The main challenge when assessing a data-driven method is the large number of degrees of freedom with respect to various aspects which include the choice of the input variables, the NN structure, the quantity and particularly the quality (representativeness) of the available data as well as their preparation and splitting between training and testing. Table 2 summarizes the main choices required

Table 2
Main decision to be taken for the data-driven approach.

Data Sampling and Splitting Sections 2.1–2.3 and 4.3	• Thermochemical state of the system • Training/ testing splitting • Training/ validation splitting • Sample weighting
Data Filtering Section 2.2	• Filter function • Filter aspect ratio • Filter sizes
Input Sections 4.2.1 and 4.2.2	• Fluid/scalar variables • Spatial variables
NN hyper-parameters Section 4.2.3	• # layers • # neurons • Activation function • Loss function

by the data-driven approach and points to the corresponding sections.

The performance of a neural network depends on a set of so-called hyper-parameters, comprising of the number of layers, number of neurons and their connectivity, choice of the activation and cost functions, and choice and settings of the optimization algorithm. The interplay between input variables and hyper-parameters is not trivial and the latter cannot be determined as part of the optimization problem during network training and need to be defined beforehand through some form of heuristics [39]. It is important to note that a different set of input variables (i.e. a different model) requires a dedicated choice of hyper-parameters to ensure a fair comparison of the input configurations (introduced in Sections 4.2.1 and 4.2.2). The main objective is to find a network topology that allows high performance, while also enabling assessment of relative input quality influence and removing potential biases arising from inappropriate hyper-parameter choice which can result in underfitting or overfitting. The optimal configuration should therefore lie between these two extreme situations: in the literature, this is often denoted as balanced model capacity [39]. The optimal topologies for every input configuration are chosen in Section 4.2.3 and the selected networks are eventually used in the final assessment where the performance of the models is compared on the test datasets.

4.2. Network input selection - model construction

Concerning the inputs to the network, three main aspects are discussed in this section: (i) the feature selection based on domain knowledge and descriptive statistics, (ii) the selection of the locations where the features are extracted, and (iii) the choice of the network topology in terms of number of neurons and hidden layers.

4.2.1. Input features

The high-fidelity simulations provide a complete description of the system and multiple quantities can potentially be used as inputs for the neural networks. A preliminary statistical screening is performed using as indicators the Pearson's correlation coefficients between the model target variable COEFF and various candidate quantities extracted from both the first and second (denoted with subscript 2) wall-adjacent cells (Fig. 5). All samples from the available datasets (equally-weighted) and time-steps were used to compute the average correlation coefficients.

As expected, ΔT (the temperature difference between the first wall-adjacent cell and the wall) has the highest correlation with COEFF and the correlation drops to almost half for the temperature in the second cell ΔT_2. The information on the fluid flow is also beneficial and the velocity component with the largest correlation is the one perpendicular to the wall. Differently from the other quantities, in this case a non-dimensionalization is proposed

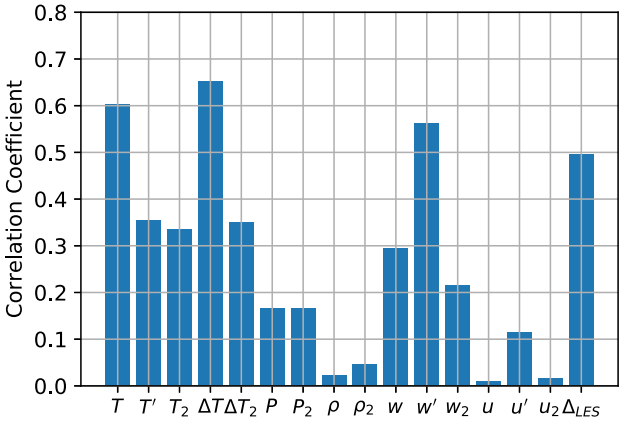

Fig. 5. Absolute value of the Pearson's correlation coefficients between the model target variable COEFF and various candidate quantities extracted from both the first and second (denoted with subscript 2) wall-adjacent cells (i' denotes the rms of quantity i; w and u refer to the velocities perpendicular and parallel to the wall, respectively).

in terms of a local, cell-based Reynolds number:

$$Re_W = \frac{w\Delta_{LES}}{\nu} \tag{4}$$

This transformation has a negligible impact on the correlation coefficient and does not modify the sign of the velocity, allowing the network to distinguish between flow impingement and ejection. In addition, Re_W allows the implicit consideration of Δ_{LES} as input. In fact, despite having a large correlation coefficient with *COEFF*, we refrained from including Δ_{LES} directly as input in dimensional form since it was found to promote overfitting in tests.

The cell pressure P shows some degree of correlation since at higher P the boundary layer tends to become thinner. Therefore, the corresponding increase in *COEFF* is indicative of the under-resolution of the boundary layer by the LES grid. The pressure is preferred over density as it shows both stronger correlations with *COEFF* and weaker correlations with the (already chosen) temperature ($\simeq 0.02$). In general, a better cumulative variance reduction is expected in a non-linear model combining variables that are weakly correlated.

Examination of RMS quantities indicates a particularly high correlation coefficient for the wall-normal velocity fluctuations, w'. This quantity indicates the presence of inhomogeneities in the vertical flow structures within the LES cell, leading to potential inversions of the local vertical flow orientation (wall impingement and ejection). High sub-grid vertical flow fluctuations are expected to increase wall heat transfer and indicate the need for substantial wall model intervention (i.e. an increase of *COEFF*). In an actual LES solver, the accuracy of modelled sub-grid fluctuations may depend considerably on closure selection and near-wall discretisation. Hence, in order to decrease the dependence on the turbulence model and constants, we refrained from using RMS quantities as inputs.

We note that while the temporal evolution of quantities could also be adopted as inputs, these are not explored in the present work.

Additional information can be obtained by linearly combining some of the quantities presented before. The results can work as a multi-linear reference for the more complex non-linear ANN models developed in the following sections. A summary of the Pearson's correlations of these multi-linear quantities with *COEFF* is presented in Table 3.

Table 3 highlights that the inclusion of the information contained in the second cell plays an important role in increasing

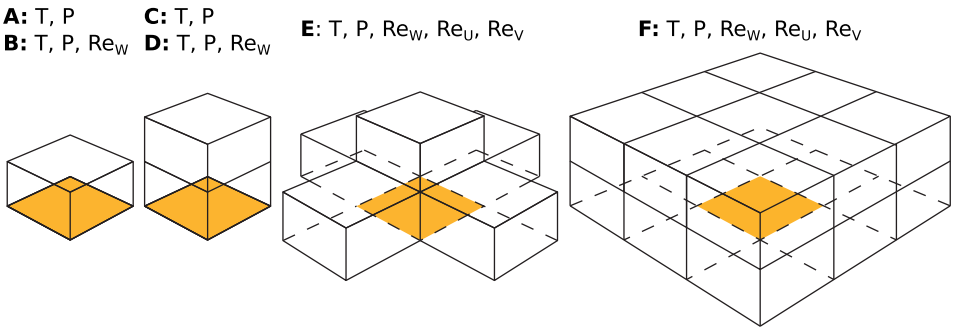

Fig. 6. Graphical representation of the six configurations in terms of input quantities (listed on top) and input locations (cells around the coloured wall surface where WHT is computed.).

Table 3
Pearson's correlation coefficients of multi-linear models.

Configuration	Quantities	Corr. Coeff
A	ΔT, P	0.63
B	ΔT, P, Re_W	0.65
C	ΔT, P, ΔT_2, P_2	0.76
D	ΔT, P, Re_W, ΔT_2, P_2, Re_{W-2}	0.77

the value of the correlation coefficient. The effect of Re_W, on the other hand, is marginal. Nevertheless, this does not necessarily mean that this quantity cannot be useful when developing neural network-based models. Due to their non-linear nature, the models may be able to exploit better the information contained in Re_W. The results in Table 3 indicate that the locations where the quantities are extracted is important and should be examined.

4.2.2. Network input configurations

Following the reasoning of the previous section, six input configurations (i.e. six distinct model variants) are considered and are denoted as A, B, C, D, E, and F. Fig. 6 shows graphically these configurations in terms of input quantities (listed on top) and input locations (cells displayed) around the coloured target wall surface at which WHT is computed.

Cases A and B make use of quantities only in the first wall-adjacent cell, while C and D employ also the second cell. Case E includes neighboring cells on the wall whereas case F includes all cells with a vertex in common with the target cell.

While T and P are always included among the inputs, the cell-based Reynolds number is included in cases B and D only in the wall-normal direction whereas for the more complex cases E and F all three components were considered since convection within the stencil is expected to provide additional structural information.

4.2.3. Network topology

Since the input and output layers have fixed dimensions, the network topology is defined by the number of hidden layers and the number of neurons per layer. To simplify the process and reduce its computational cost (i.e. a reduced number of combinations), all hidden layers share the same number of neurons n. To avoid networks that are strongly asymmetric, n is related to the number of inputs N_{inp} (the ratio n/N_{inp} is varied between 1/6 and roughly 4) while the number of hidden layers is varied between 1 and 6. A total of 80 distinct combinations of hyper-parameters have been tested for each one of the six input configurations.

4.3. Network training - model tuning

4.3.1. Data pre-processing

The following operations are performed to prepare the data for training:

- **Data Augmentation:** the original dataset is expanded performing four 90-degree rotations of the frame of reference around the wall-normal axis. This process reduces the risk of finding spurious relations between the features and the target, which is the case in the presence of a significant tumble flow inside the cylinder like in the TUD engine setup.
- **Data Splitting:** in addition to the splitting proposed in Section 2.3, 20% of the training samples are randomly extracted to construct a validation set. This set is used to make decisions at a higher level, e.g. to select the best topology and it is not used directly to tune the network weights.
- **Data Normalization:** the input features are normalized to enforce zero mean and unit variance. This process can improve and accelerate the training process.
- **Data Weighting:** for a given wall surface area, filtering with a smaller filter size generates more filtered samples. The difference in sample counts has been compensated with weights so that every filter size has the same importance. The same argument applies for differences in number of samples between datasets. The weight associated with a sample i is therefore:

$$W_i = \frac{Z_{dataset}}{Z_{filter}} \cdot \frac{Z_{all}}{Z_{dataset}} \cdot F$$

where $Z_{dataset}$ is the number of samples of the dataset to which i belongs, Z_{filter} is the number of samples of the same dataset and with the same filter size as i, Z_{all} is the total number of samples across all the datasets and filter sizes, and F is a user-defined weighting factor which is always equal to one except for the TCC samples where $F = 2$. This choice is made to weight the TCC dataset more heavily than the Syngas one during training, since it is closer to the targeted IC engine applications.

4.3.2. Training process

The training process is mainly determined by the number of "epochs", where the completion of one epoch means using each of the training samples once during the optimization. An adaptive strategy is used in this work, where new epochs are progressively added only if they are beneficial for the validation score. During each epoch, a mini-batch approach with a 2048 samples is used. Each network is initialized with random weights and the training is repeated multiple times: 1000 training realizations for configurations A-B-C-D and 10 for E-F. Particularly for smaller networks, this large number of realizations was necessary to assure statistical convergence of the first two moments of the R2 score for every hyper-parameter setting. Note that 1000 training realizations result in an equal number of variants of network weights so that the performance during testing is assessed for all and the results are shown in terms of ensemble statistics.

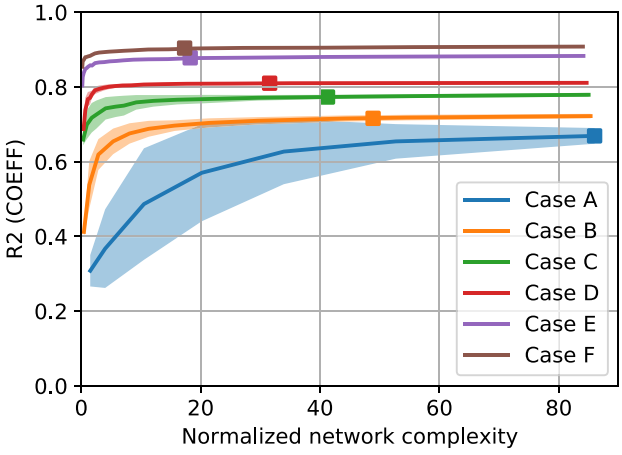

Fig. 7. R2 score of COEFF versus normalized network complexity for the validation dataset for the different sets of input features. Lines and shaded areas represent ensemble average and standard deviation, respectively. The architectures selected for each case are highlighted by a bold square.

4.3.3. Network choice

A specific challenge in the visualization and synthesis of the results is the wide range of network variants and sizes. The variations in input features and their locations are summarized by the case name (A to F). The 80 case-specific network variations are lumped into a single representative coordinate, the normalized network complexity C:

$$C = \left(N_{inp} \cdot n + N_{out} \cdot n + n^2 \cdot (n_{hl} - 1)\right)/N_{inp}^2 \tag{5}$$

defined as the ratio between the total number of weights in the network and the number of inputs squared. Here, N_{inp} and N_{out} are the number of inputs and outputs, respectively; n is the number of neurons per layer and n_{hl} is the number of hidden layers. This coordinate allows for a compact visualization of all cases and provides a qualitative measure relating the number of inputs and the number of network weights.

The best topologies are chosen with the help of the R2 score of the target variable *COEFF*, which is commonly used in literature as key performance indicator (KPI) for this kind of problems. It is a measure of the variance of a target quantity captured by the model. Fig. 7 shows R2 as a function of the normalized network complexity for all configurations. The R2 scores of *COEFF* are computed using the validation data, which consists of the 20% of random samples excluded from training. The shaded areas represent the standard deviation computed over the different training realizations.

Overall, a monotonic increase in R2 is observed when more input variables are considered (from Case A to Case F). An increased network complexity (i.e. increase of the number of weights) results in a progressive saturation of the R2 score increment. Despite the large variation in network complexity, there is no sign of over-fitting for the validation dataset in the form of a performance degradation for larger networks.

Underfitting is observed for low-complexity networks, indicating that such networks are unable to fully exploit the information contained in the data. In addition, these networks are subjected to a large sensitivity with respect to the weight initialization as visible from the large R2 variance. This is especially the case for configurations that use a small number of input features (notably case A). The high sensitivity for case A indicates the presence of a large number of local minima in the cost function. In particular, the absence of any information from the second wall-adjacent cell seems to hinder the model from acquiring a robust signal recognizing the

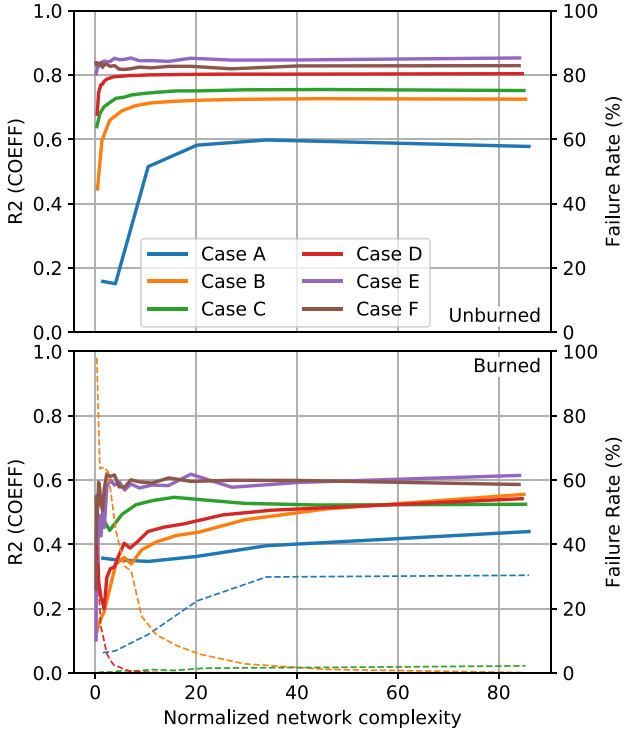

Fig. 8. Ensemble average of R2 score of COEFF (solid lines) and failure rate (dashed lines) versus normalized network complexity for the *Unburned* (top) and *Burned* (bottom) testing datasets.

different states of the system near the wall. While such variability is reduced as the complexity increases, even the largest complexity exhibits a moderate level of variance. The aforementioned issue is attenuated for configuration B where the inclusion of Re_W appears to help the model to indirectly reconstruct the information from the second cell.

The optimal topology is the one that combines, at the lowest complexity, good performance, independence from the initial NN weights and stability with respect to the complexity itself. The selected topologies that match the aforementioned criteria are highlighted with bold squares in Fig. 7 and summarized in Table 4.

4.4. Network testing - model performance assessment

Network testing is performed separately for the three sample clusters “unburned”, “flame boundary” and “burned” defined in Section 2.4. It is worth repeating that testing and training datasets are completely separated. Note that the samples in the “flame boundary” deserve a different assessment and will be presented in Fig. 12.

We start with a comparison of the average R2 scores for the “unburned” and “burned” (upper and lower plot, respectively) test cases as shown in Fig. 8.

For the “unburned” case (TUD-NR dataset), similar results as for the validation dataset (Fig. 7) are observed. The performance of networks B-F is more clustered. A mild manifestation of overfitting is observed for case A, where a slight deterioration is noted with increasing complexity. Overall, a similar performance between TCC and TUD-NR is expected since both cases are non-reactive and with comparable ranges in engine Reynolds number and local wall heat fluxes.

Conversely, such a similarity does not exist for the “burned” test case (TUD-R dataset), where close to one order of magnitude higher wall heat fluxes are encountered due to the high temper-

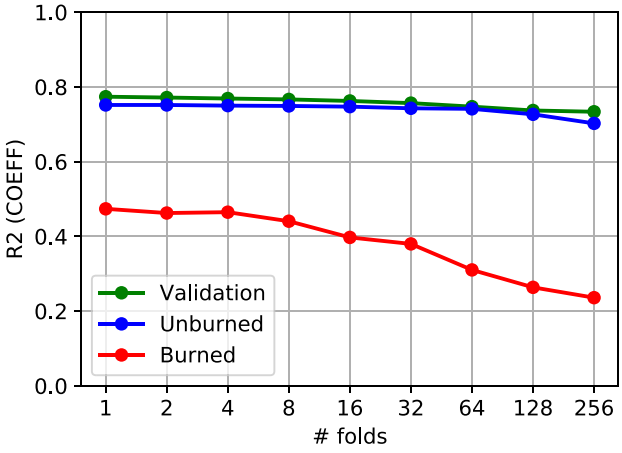

Fig. 9. R2 score as a function of the number of folds for case C. Each point is obtained as the ensemble average of the R2 scores over the range of normalized network complexities 30–60.

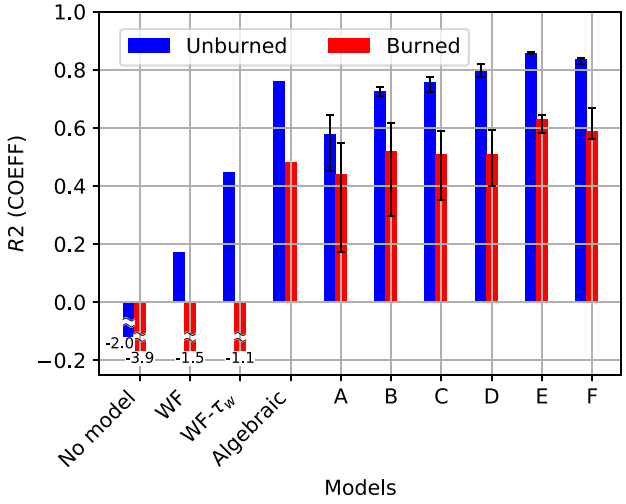

Fig. 10. Absolute R2 scores for the different models for the *Unburned* (blue) and *Burned* (red) testing datasets. For the neural network models, results are drawn in terms of median (bar) and 10–90 percentiles (error bars) using the network hyperparameters of Table 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

atures originating from combustion. This makes the "burned" test case more challenging. For some of the test realizations for a given network topology, negative R2 scores were obtained. Since the R2 score is bounded on the positive side (to unity) but not on the negative side, a few badly-performing realizations can have a detrimental influence on the mean value. Quantitative information of negative R2 scores is not relevant and can be simply categorized as bad performance. In order to dismiss the ensuing quantitative bias while still properly representing the occurrence rate of poor performance, we opted for a representation of the results in terms of "failure rate", defined as the percentage of realizations with a negative R2 score, in combination with the average R2 score, now depicting only those cases with a positive score. Note that the failure rate for the "unburned" test case is always zero.

For the "burned" region, the models show a non-zero failure rate for the lower complexity. However, scores in most cases become positive at higher complexity. The only exception to this trend is for configuration A, where the failure rate is always above 10% and increases to roughly 30% with increasing complexity. This counter-intuitive behaviour is related to the aforementioned over-fitting of the training data obtained for this particular network configuration with the largest complexities. Overall, R2 scores decrease significantly for all models and there is a convergence of performance between cases B, C and D and a mild dependency on the network complexity.

The absence of over-fitting for the configurations B to F using an 80% training 20% validation split is an indication that the training dataset is statistically sufficiently homogeneous combined with a sample overhead with statistically equivalent samples. This means that retaining 20% of the samples for validation results in almost the same statistical structure and makes over-fitting of the network difficult to observe. To confirm the training sample overhead and the observation of a plateauing behaviour, we varied the number of samples used for training for case C. In statistical modelling, this is referred to as n-fold cross-validation, and is typically done to assure statistical relevance of the data and the susceptibility of the model on input data. Here, we perform the training-validation procedure for case C using only a subset of the full training datasets. We split the original training dataset by a factor $n = 1$, 2, 4, 8, 16, 32, 64, 128, 256, and subsequently apply the 80%/20% split of the subsets. Fig. 9 shows the ensemble average of R2 score over the range of normalized network complexities 30–60 for every subset of data used for training for case C. A reduction of the subset used for training has a minor influence on the performance

Table 4
Architectures of the selected networks.

Configuration	#Inputs	#Hidden layers	#Neurons
A	2	6	8
B	3	5	10
C	4	6	11
D	6	5	16
E	30	4	69
F	90	3	243

of the validation and "unburned" test cases. For the "burned" test case, a significant breakdown in performance due to over-fitting is observed by a reduction of samples by a factor of more than 8.

An additional test was performed using only the TCC data during training for case C and for the different network hyperparameters. The results obtained (not shown) led to two main observations: first, for the "unburned" test case, larger networks showed a clear evidence of over-fitting in the form of a performance reduction; second, the absence of reactive samples in the training resulted into negative R2 scores for the "burned" test case for all complexities because the neural network focused on the non-reactive situations. These two observations confirm the importance of including reactive data in the training process.

As a next step, a single normalized complexity marked by the squares in Fig. 7 for each case is selected. Fig.10 compares the R2 scores for the "unburned" and "burned" cases with the reference models introduced in Section 3.2. The "flame boundary" zone is excluded from this representation since it is characterized by a 100% failure rate. For the neural network models, the results are plotted in terms of the 10-50-90 quantiles to give an idea of the actual distribution of the scores, which are obtained after excluding the failed models ($R2 < 0$). For the reference models, negative R2 scores have been clipped and the effective numbers are indicated in the plot.

Starting from the "unburned" test case (blue bars), the "no model" method scores a negative R2 value, emphasizing the evident need for a model. The standard wall function expectedly shows some improvement. When considering the artificially enhanced $WF - \tau_W$ variant (wall shear stress fed from filtered DNS data), it is observed that the error source from the wall shear

stress model can account for one third of the error in *COEFF*. The algebraic model developed in Bolla et al. [27] shows a significant improvement and is performing similarly to the NN in case C. Under reactive conditions (red bars), both wall function variants show some improvement compared to the "no model" approach, but the absolute performance deteriorates significantly. The combustion-induced near-wall laminarization (due to high temperature and hence viscosity) along with other non-equilibrium terms such as pressure-work and chemical reaction results in conditions that are exceedingly removed from those for which conventional (turbulent flow) wall models have been developed. Hence, it is perhaps not surprising that the correct wall shear stress inputs do not provide much assistance for such models. The algebraic model continues to exhibit performance comparable to the neural network models. The "burned" case shows a more leveled performance along models at lower R2 values and higher levels of variabilities between realizations.

Among the developed models, model A displays the lowest score and the largest variability. Models B, C and D have quite similar scores with the latter having slightly better performance in the "unburned" case and a slightly smaller variability. Model E performs best overall, combining the largest scores and smallest variability in both test sets. All neural network models perform better than the wall function-based models examined in this work.

It is interesting to note that the algebraic model and configuration C use the same input features and have similar performance. This indicates that the features in the algebraic model effectively extract all of the important information related to heat flux behaviour. In fact, the neural network that can be seen as an advanced regressor and does not provide any remarkable improvement.

The investigations so far provide a collective view of functionality over different grid sizes. To obtain additional detail, the model performance is assessed in the following as a function of the dimensional filter size. As this is a model-centric study, we want to discern the contributions of the various models compared to the "no model" method, and introduce the relative R2 score:

$$R2_{rel} = 1 - \frac{\sum_i (y_{i,-DNS} - y_{i,-model})^2}{\sum_i (y_{i,-DNS} - y_{i,-nomodel})^2} \tag{6}$$

The difference with the standard definition is in the normalization which is now related to the performance of the "no model" case. This formulation allows for a fairer comparison between the scores computed for different datasets. In fact, the numerator and the denominator are expected to have a similar trend when the dataset changes, reducing the sensitivity of $R2_{rel}$ on the choice of the dataset itself. $R2_{rel}$ is equal to 0 if performance is the same as that of "no model", while it becomes equal to 1 for a perfect model.

The $R2_{rel}$ scores as functions of the filter size for all models for the "unburned" (upper) and "burned" (lower) datasets are presented in Fig. 11. All models follow approximately the same trend: the lowest values are obtained for the smallest Δ_{LES} due to the improvement in the "no model" performances; then the $R2_{rel}$ scores increase approaching a plateau, showing low sensitivity on Δ_{LES}. The plateau means that model performance is deteriorating at a similar rate to the "no model" approximation. As the "no model" grows less accurate with increasing wall distances, a constant $R2_{rel}$ score means that smaller filter widths provide better (absolute) modelling results. Again, all NN models perform better than the classical WFs while the algebraic model performs comparably with configuration C.

For the "flame boundary" region, the performance of the models is significantly worse ($R2 < 0$). As the $R2$ score is not bounded in the negative direction, it is not a suitable metric in this case and is therefore replaced with the ratio between the average heat

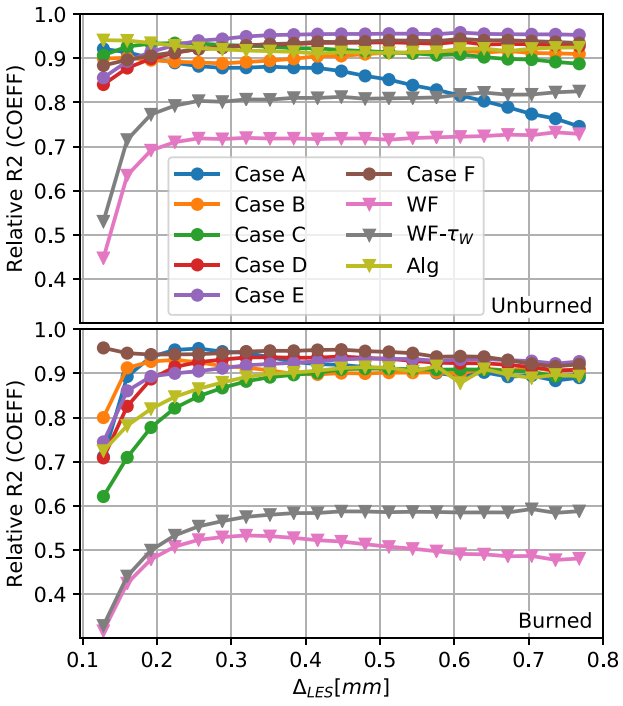

Fig. 11. Relative R2 scores versus dimensional filter size (Δ_{LES}) for the different models for the *Unburned* (blue) and *Burned* (red) testing datasets using the network hyper-parameters indicated in Table 4. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

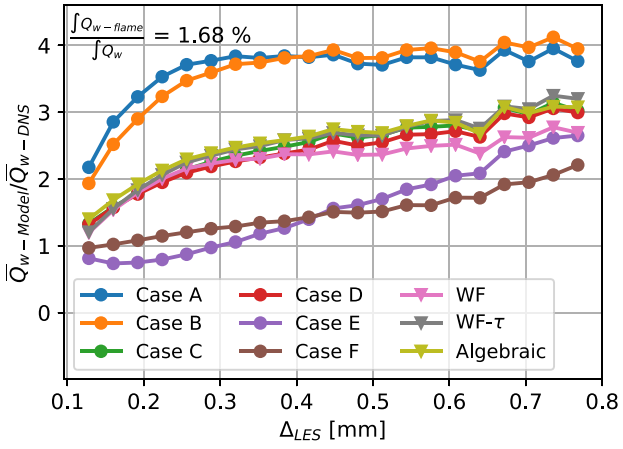

Fig. 12. Ratio between average modelled and filtered DNS heat fluxes versus dimensional filter size (Δ_{LES}) for the different models for the *Flame boundary* testing dataset using the network hyper-parameters indicated in Table 4.

fluxes (Fig. 12), which highlights the inherent behaviour discussed in Section 2.4 with the modelled heat flux exceeding the DNS one. In general, all models perform quite poorly in this region with configurations E and F having the best behaviour. Nevertheless, the contribution of this region to the overall heat flux is minor (less than 2%), discouraging improvement attempts that risk functionality in "unburned" and "burned" samples.

A final stress test on the model is performed by changing the cell aspect ratio. As described in Section 2.2, a 1:2 aspect ratio was used during training. In actual WMLES work, however, the aspect ratio can vary significantly depending on (i) the physics and level of fidelity that one is trying to reproduce, (ii) wall geometry, (iii) computational power availability, and importantly, (iv) the

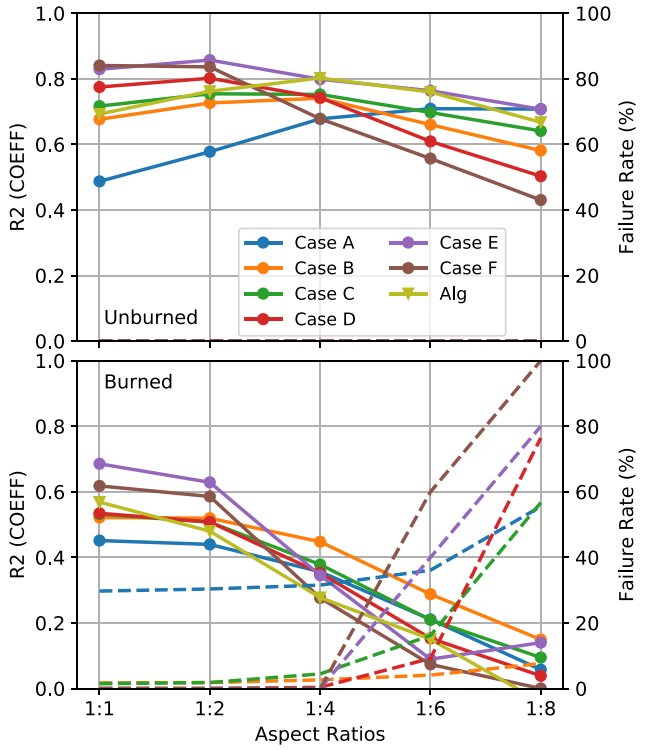

Fig. 13. Ensemble average of R2 score (solid lines) and failure rate (dashed lines) versus filter aspect ratio for the *Unburned* (top) and *Burned* (bottom) testing datasets. Models have been trained exclusively with filter aspect ratio 1:2. Network hyper-parameters indicated in Table 4.

meshing structure. A situation where the aspect ratio is constant for all near-wall cells is therefore a significant simplification. The test is carried out using aspect ratios of 1:1; 1:2; 1:4; 1:6 and 1:8, which in practice refers to a variation of the filtering process. The results are presented (only for the selected complexities in Fig. 7) in terms of average R2 score and failure rate separately for the "unburned" and the "burned" regions in Fig. 13.

It is clearly noticeable that larger input stencils lead to higher sensitivity to the filter aspect ratio. Especially in the reactive case, the behaviour of the networks E and F is not robust, with a degradation that approaches 100% failure rate for the aspect ratio 1:8. A physical reasoning to this observation stems from the fact that the training and testing stencils capture strongly contrasting physical scales of the near-wall flow, potentially leading to misinterpretation by the trained model. The network that performs best in this case is the network C, with networks B and D also performing moderately well. From this analysis, it is clear that the size of the input stencil does not directly correlate with model performance.

It can be argued that, for cases E and F, the slightly larger scores in the test sets do not compensate for (i) the increase in complexity for their implementation and (ii) their larger sensitivity to different grid configurations. These more complex stencils also require a higher quality mesh, discouraging their use in complex geometries. The most promising configurations, combining accuracy and robustness, appear to include B, C and D. As their complexity in both training and LES solver implementation is similar, they can be compared in terms of the quantitative scores in the present work. In this respect, network D seems to be the optimal trade-off solution that combines relatively high scores with small sensitivity to the initial conditions and to the grid aspect ratio. In addition, network D has reasonable performance when applied to the "flame boundary" region.

Due to the nature of the present *a priori* study, we do not report the added computational cost of wall modelling approaches. The added computational cost of wall modelling in WMLES is often reported to be small in comparison to the remainder of the flow solution [8]. A computationally intensive full non-equilibrium turbulent boundary layer equations (TBLE) wall model has been reported to induce an extra cost of 100–150% [40]. While significant, even such increases are thought to be small in comparison to the savings afforded by wall modelling [8]. In contrast to such approaches with operations including matrix inversion [40], we note that the presently tested models and ANNs represent a series of relatively straightforward algebraic operations.

5. Conclusions

In the present study, data-driven wall heat transfer models for WMLES have been developed using neural networks trained and tested on various sets of high-fidelity simulation data, both in real IC engines and engine-like configurations. Stressing the importance of proper data selection and formulation for the successful generation of models, the present work aimed at evaluating the criteria for successful model generation. Observations were divided into three separate clusters denoted as "unburned", "burned", and "flame boundary", each displaying distinct characteristics and a varying degree of difficulty for model generation. Variations of (i) input physical quantity, (ii) input data stencil type and (iii) stencil dimensions allowed for a unique cross-examination of the selections.

The following conclusions are drawn from this work:

- Neural networks can be used to successfully reconstruct the complex unknown relation between features at LES resolution and *COEFF* or, indirectly, the heat flux, with good performance.
- With a computational cost similar to the reference algebraic models for the online evaluation during an LES simulation, the neural network-based models show improved performances and represent a valid alternative to the algebraic heat flux models currently available for LES.
- Among the investigated configurations, case D is identified as the best compromise between performances and complexity. The information contained in Re_w and in the second wall-normal cell are beneficial for the model accuracy. Simpler configurations have worse performances with a similar computational cost while the additional complexity associated to the configurations E and F is not compensated by sufficient improvement in the model accuracy.

While the present work demonstrates the feasibility of functional ANN-based WHT models, it does not claim to resolve the challenges of WHT modelling in general. In the context of future works, various aspects - beside an a posteriori model assessment - could be considered for further developments. Here we mention three of them. (i) The techniques presented here could be more universally applicable if conditions were expanded to a wider span of engine conditions, particularly in terms of Reynolds number. (ii) The inclusion of a temporal dependence in the model could be beneficial and recurrent neural networks (RNN) seem a promising candidate for capturing time-dependent patterns. (iii) In the detection of spatial structures and relationships in complex data, convolutional neural networks (CNNs) are an established approach that can be explored.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Matteo Impagnatiello: Conceptualization, Methodology, Investigation, Writing – original draft. **Michele Bolla:** Conceptualization, Supervision, Writing – original draft. **Karri Keskinen:** Methodology, Writing – review & editing. **George Giannakopoulos:** Data curation, Visualization, Software. **Christos E. Frouzakis:** Resources, Writing – review & editing. **Yuri M. Wright:** Resources, Writing – review & editing, Project administration. **Konstantinos Boulouchos:** Funding acquisition, Supervision.

Acknowledgements

This research used resources of the Argonne Leadership Computing Facility under Contract DE-AC02-06CH11357 and the Swiss National Supercomputing Center (CSCS) under project ID 753. Financial support from the Forschungsvereinigung Verbrennungskraftmaschinen (FVV, project no. 1286: “Wall heat transfer processes in spark ignition engines”), the Swiss Federal Office of Energy (Grants SI/501615-01, SI/501584-01 and SI/501931-01) and the Swiss Competence Center for Energy Research (SCCER Mobility) is gratefully acknowledged.

References

[1] J.N. Kutz, Deep learning in fluid dynamics, J. Fluid Mech. 814 (2017) 1–4.
[2] S. Brunton, B. Noack, P. Koumoutsakos, Machine learning for fluid mechanics, Annu. Rev. Fluid Mech. 52 (2020), doi:10.1146/annurev-fluid-010719-060214.
[3] Z. Wang, K. Luo, D. Li, J. Tan, J. Fan, Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation, Phys. Fluids 30 (2018) 125101, doi:10.1063/1.5054835.
[4] C.J. Lapeyre, A. Misdariis, N. Cazard, D. Veynante, T. Poinsot, Training convolutional neural networks to estimate turbulent sub-grid scale reaction rates, Combust. Flame 203 (2019) 255–264.
[5] M. Bode, M. Gauding, Z. Lian, D. Denker, M. Davidovic, K. Kleinheinz, J. Jitsev, H. Pitsch, Using physics-informed super-resolution generative adversarial networks for subgrid modeling in turbulent reactive flows, arXiv preprint arXiv:1911.11380 (2019).
[6] R. Maulik, O. San, A. Rasheed, P. Vedula, Data-driven deconvolution for large eddy simulations of Kraichnan turbulence, Phys. Fluids 30 (2018) 125109, doi:10.1063/1.5079582.
[7] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. 378 (2019) 686–707, doi:10.1016/j.jcp.2018.10.045.
[8] S. Bose, G. Park, Wall-modeled large-eddy simulation for complex turbulent flows, Annu. Rev. Fluid Mech. 50 (2018) 535–561, doi:10.1146/annurev-fluid-122316-045241.
[9] K. Suga, T. Sakamoto, Y. Kuwata, Algebraic non-equilibrium wall-stress modeling for large eddy simulation based on analytical integration of the thin boundary-layer equation, Phys. Fluids 31 (2019) 075109, doi:10.1063/1.5099658.
[10] H.J. Bae, A. Lozano-Durn, S. Bose, P. Moin, Dynamic slip wall model for large-eddy simulation, J. Fluid Mech. 859 (2019) 400–432, doi:10.1017/jfm.2018.838.
[11] C. Jainski, L. Lu, A. Dreizler, V. Sick, High-speed micro particle image velocimetry studies of boundary-layer flows in a direct-injection engine, Int. J. Engine Res. 14 (2013) 247–259, doi:10.1177/1468087412455746.
[12] C.-P. Ding, B. Böhm, A. Renaud, A. Dreizler, S. Jakirli, Experimental characterization of the velocity boundary layer in a motored ic engine, Int. J. Heat Fluid Flow 71 (2018), doi:10.1016/j.ijheatfluidflow.2018.04.014.
[13] M. Schmitt, C. Frouzakis, Y. Wright, A. Tomboulides, K. Boulouchos, Direct numerical simulation of the compression stroke under engine-relevant conditions: evolution of the velocity and thermal boundary layers, Int. J. Heat Mass Transf. 91 (2015) 948–960, doi:10.1016/j.ijheatmasstransfer.2015.08.031.
[14] P. Ma, M. Greene, V. Sick, M. Ihme, Non-equilibrium wall-modeling for internal combustion engine simulations with wall heat transfer, Int. J. Engine Res. 18 (2017), doi:10.1177/1468087416686699.
[15] K. Keskinen, J. Koch, Y. Wright, M. Schmitt, M. Nuutinen, O. Kaario, V. Vuorinen, M. Larmi, K. Boulouchos, Numerical assessment of wall modelling approaches in scale-resolving in-cylinder simulations, Int. J. Heat Fluid Flow 74 (2018) 154–172, doi:10.1016/j.ijheatfluidflow.2018.09.016.
[16] K. Truffin, C. Angelberger, S. Richard, C. Pera, Using large-eddy simulation and multivariate analysis to understand the sources of combustion cyclic variability in a spark-ignition engine, Combust. Flame 162 (2015), doi:10.1016/j.combustflame.2015.07.003.
[17] A. Misdariis, O. Vermorel, T. Poinsot, LES of knocking in engines using dual heat transfer and two-step reduced schemes, Combust. Flame 162 (2015), doi:10.1016/j.combustflame.2015.07.023.
[18] P. Janas, I. Wlokas, B. Bhm, A. Kempf, On the evolution of the flow field in a spark ignition engine, Flow, Turbul. Combust. 98 (2016), doi:10.1007/s10494-016-9744-3.
[19] C. He, W. Leudesdorff, F. Mare, A. Sadiki, J. Janicka, Analysis of in-cylinder flow field anisotropy in ic engine using large eddy simulation, Flow, Turbul. Combust. 99 (2017), doi:10.1007/s10494-017-9812-3.
[20] G.K. Giannakopoulos, K. Keskinen, J. Koch, C.E. Frouzakis, Y.M. Wright, K. Boulouchos, M. Schmidt, C.-P. Ding, B. Böhm, A. Dreizler, Characterising the evolution of boundary layers in ic engines by combined laser-optical diagnostics, direct numerical and large-eddy simulations, Flow, Turbul. Combust., (2021) under review.
[21] G.K. Giannakopoulos, C.E. Frouzakis, P.F. Fischer, A.G. Tomboulides, K. Boulouchos, LES of the gas-exchange process inside an internal combustion engine using a high-order method, Flow, Turbul. Combust. 104 (2) (2020) 673–692.
[22] C. Mandanis, M. Schmitt, J. Koch, Y. Wright, K. Boulouchos, Wall heat flux and thermal stratification investigations during the compression stroke of an engine-like geometry: acomparison between LES and DNS, Flow, Turbul. Combust. 100 (2017), doi:10.1007/s10494-017-9879-x.
[23] W. Leudesdorff, T. Unger, J. Janicka, C. Hasse, Scale-resolving simulations for combustion process development, MTZ Worldwide 80 (2) (2019) 62–67.
[24] X. Yang, S. Zafar, J.-X. Wang, H. Xiao, Predictive large-eddy-simulation wall modeling via physics-informed neural networks, Phys. Rev. Fluids 4 (3) (2019) 034602.
[25] X.L.D. Huang, X.I.A. Yang, R.F. Kunz, Wall-modeled large-eddy simulations of spanwise rotating turbulent channels comparing a physics-based approach and a data-based approach, Phys. Fluids 31 (12) (2019) 125105, doi:10.1063/1.5129178.
[26] J. Kim, C. Lee, Prediction of turbulent heat transfer using convolutional neural networks, J. Fluid Mech. 882 (2020) A18.
[27] M. Bolla, M. Impagnatiello, K. Keskinen, G.K. Giannakopoulos, C.E. Frouzakis, Y.M. Wright, K. Boulouchos, Development of an algebraic wall heat transfer model for LES in IC engines using DNS data, Proc. Combust. Inst. 38 (2021) 5811–5819.
[28] Nek5000 version v17.0. Argonne National Laboratory, IL, U.S.A., URL https://nek5000.mcs.anl.gov.
[29] P. Schiffmann, S. Gupta, D. Reuss, V. Sick, X. Yang, T.-W. Kuo, TCC-III engine benchmark for large-eddy simulation of ic engine flows, Oil Gas Sci. Technol. 71 (1) (2016) 3.
[30] M. Jafargholi, Direct Numerical Simulation of Turbulent Premixed Syngas Flames in Closed Domains, ETH Zurich, 2018 Ph.D. thesis.
[31] J. Koch, S. Geringer, D. Farrace, S. Pandurangi, M. Bolla, Y.M. Wright, M. Jafargholi, C. Frouzakis, K. Boulouchos, Assessment of two premixed LES combustion models in an engine-like geometry, SAE Technical Paper, SAE International, 2018, doi:10.4271/2018-01-0176.
[32] B. Franzelli, E. Riber, L.Y. Gicquel, T. Poinsot, Large eddy simulation of combustion instabilities in a lean partially premixed swirled flame, Combust. Flame 159 (2) (2012) 621–637.
[33] G. Smith, D. Golden, M. Frenklach, N. Moriarty, B. Eiteneer, M. Goldenberg, C. Bowman, R. Hanson, S. Song, W. Gardiner, V. Lissianski, Z. Qin URL http://combustion.berkeley.edu/gri-mech/version30/text30.html.
[34] A. Dreizler, B. Böhm, Advanced laser diagnostics for an improved understanding of premixed flame-wall interactions, Proc. Combust. Inst. 35 (2014), doi:10.1016/j.proci.2014.08.014.
[35] C. Plengsaard, C. Rutland, Improved engine wall models for large eddy simulation (LES), SAE Technical Papers, 2, 2013, doi:10.4271/2013-01-1097.
[36] C. Angelberger, T. Poinsot, B. Delhay, Improving near-wall combustion and wall heat transfer modeling in SI engine computations, SAE Technical Paper, SAE International, 1997, doi:10.4271/972881.
[37] C. Rakopoulos, G. Kosmadakis, E. Pariotis, Critical evaluation of current heat transfer models used in CFD in-cylinder engine simulations and establishment of a comprehensive wall-function formulation, Appl. Energy 87 (2010) 1612–1630, doi:10.1016/j.apenergy.2009.09.029.
[38] D. Kingma, J. Ba, Adam: a method for stochastic optimization, in: International Conference on Learning Representations, 2014.
[39] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.
[40] G.I. Park, P. Moin, Numerical aspects and implementation of a two-layer zonal wall model for LES of compressible turbulent flows on unstructured meshes, J. Comput. Phys. 305 (2016) 589–603.