

# Data-Driven Combustion Mode Identification for MILD Combustion in Coarse Grained Field

Group A05

C. Alcántara Collado, D. Das, G.M.J. Gelissen, B. Gomez-Guillamon Jurado, M. González Piñeiro,  
N.A. Keating, S. Ramasubramanian, O. Ruan, J. de Steenhuijsen Piters, A. Trigueros Alonso,

Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, HS Delft 2629, Netherlands

## Abstract

Moderate or Intense Low-oxygen Dilution (MILD) combustion has recently emerged as a low-emission combustion method, capturing the interest of field experts. However, MILD combustion features two distinct modes within the combustion field, presenting modeling challenges for Large Eddy Simulation (LES) that could offer a favorable balance between cost and accuracy. In this paper, a multi-perceptron Neural Network (NN) model is proposed for the identification of combustion modes, characterised by temperature and reaction rates, within MILD combustion under premixed conditions. The aim is to investigate and improve the usage of an existing neural network for this identification, potentially being integrated into an LES model to enhance accuracy. Gaussian filters are applied to DNS data to emulate LES data which will be inputted into the NN. Furthermore, a scalar data field is obtained through the volume fraction of local combustion mode (ignition/interaction) definition, leading to a comparison of results obtained from the two methods. Additionally, the predictive efficacy of a "zeroth-order" model is compared to the NN-based prediction. The results show a strong correlation between the predictions made by the NN and the direct outcome from the DNS data for all filter sizes. Contrarily, the zeroth-order model shows an unsatisfactory performance, especially for larger filter sizes. The results suggest that neural networks can be a viable solution for solving combustion mode identification for MILD combustion applications. A broader dataset and additional data processing are recommended to obtain a more accurate and less under-fitted model solution. The revised model shows significant improvement over the original model, suggesting the efficacy of the suggested adjustments.

**Keywords:** MILD combustion; Combustion mode; Neural Network (NN); Machine learning; Direct Numerical Simulation (DNS); Large Eddy Simulation (LES); Subgrid-scale (SGS).

## Nomenclature

### Abbreviations

CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
DNS	Direct Numerical Simulation
EGR	Exhaust Gas Recirculation
LES	Large Eddy Simulation
MILD	Moderate or Intense Low-oxygen Dilution
MSE	Mean Squared Error
NN	Neural Network
NOx	Nitric Oxides
PCA	Principal Component Analysis
SGS	Subgrid-Scale

### Greek Symbols

$\delta_{th}$	Laminar flame thickness [mm]
$\varepsilon_{MSE}$	Mean-squared error
$\omega_{CT}$	Reaction rate [ $mol \cdot L^{-1} \cdot s^{-1}$ ]
$\rho_r$	Reactant gas density [ $kg \cdot m^{-3}$ ]
$\sigma$	Standard deviation [mm]
$\Theta$	Principal gradient direction angle [deg]

### Roman Symbols

$c_p$	Specific heat capacity [ $J \cdot kg^{-1} \cdot K^{-1}$ ]
$\delta_x$	Mesh width [mm]
$k$	Filter width / Kernel size [mm]
$l_0$	Integral length scale [mm]
$\nabla c_T$	Normalised temperature gradient [ $m^{-1}$ ]
$\nabla T$	Temperature gradient [ $K \cdot m^{-1}$ ]
$\Phi_{0th}$	Zeroth-order volume fraction of local combustion mode
$\Phi_{NN}$	Neural network volume fraction of local combustion mode
$\Phi_{res}$	Fully-resolved volume fraction of local combustion mode
$Q$	Heat release rate [W]
$r_P$	Pearson correlation coefficient
$S_L$	Laminar flame speed [ $m \cdot s^{-1}$ ]
$T_b$	Burnt temperature [K]
$T_u$	Reactant temperature [K]
$u'$	Turbulence intensity [ $m \cdot s^{-1}$ ]
<i>Superscripts</i>	
$(\cdot)$	Filtered quantity
$(\cdot)^+$	Non-dimensionalised quantity
$(\cdot)^*$	Normalised quantity
$\tilde{(\cdot)}$	Favre-filtered quantity

## I. Introduction

WITHIN the realm of combustion studies, Moderate or Intense Low-oxygen Dilution (MILD) combustion technology emerges as a promising technology for improving combustion efficiency while significantly reducing pollutant emissions [1]. MILD combustion is characterised by a unique combustion process where the temperature of the incoming reactant gases surpasses the self-ignition temperature of the mixture, but remains below the temperature at which flamefront propagation occurs [2].

This technology is often implemented using an Exhaust Gas Recirculation (EGR) technique, where exhaust gases are fed back into the combustion chamber, reducing the oxygen and fuel concentrations [2, 3]. By maintaining lower maximum temperatures and minimising hot-spot formation, MILD combustion effectively mitigates Nitric Oxide (NOx) emissions and enhances thermal efficiency, offering advantages for industrial furnace systems and gas turbine plants [1, 4–6].

Experimental and computational studies have revealed the presence and coexistence of two distinct combustion modes within MILD combustion [7]. The first mode, ignition/interaction, is characterised by a uniform temperature distribution and high turbulence levels [4]. The second mode, propagation, is governed by high reaction rates and non-uniform temperature distribution, where the flame front propagates in space [8]. Since these two modes can coexist, variable  $\Phi$  is defined to represent the volume fraction of local combustion mode (ignition/interaction) [9]. Precise modelling of the combustion process requires accurate identification of these combustion modes [2, 10]. Since their discovery, extensive research has been conducted on various methods for identifying these modes [11–14].

With the advancement of Computational Fluid Dynamics (CFD) tools and increased computational capabilities, Large Eddy Simulations (LES) has emerged as a viable method for analysing and designing MILD combustion systems [14, 15]. The effectiveness of LES in this context relies heavily on the accurate modelling of Subgrid-Scale (SGS) phenomena, a task complicated by the coexistence of distinct combustion modes within MILD combustion fields [2]. Identifying these modes accurately in LES contexts is crucial for developing appropriate SGS combustion models, enhancing the precision of LES in capturing the characteristics of MILD combustion [9].

Recent studies indicate that the integration of artificial intelligence, specifically Neural Networks (NN), into combustion mode identification leads to promising results [9, 16]. With the use of Direct Numeri-

cal Simulation (DNS) data for training, emulating LES data through filtering, NNs serve as a potential tool for complementing LES computations [9, 15, 16]. However, there is a limited amount of research on their possible applications in MILD combustion mode identification.

The aim of this study is to investigate and improve the usage of a neural network for MILD combustion mode identification under premixed conditions, building upon the work of Jigjid et al. [9]. To achieve this goal, DNS data, specifically reaction rates and temperature gradient scalar coordinates fields, obtained in a study by Minamoto et al. [17] will be processed and filtered. The performance of the NN will be benchmarked using a filtered scalar field derived from DNS data, serving as the target for model predictions. Accuracy metrics obtained from these predictions will be compared to those from a basic zeroth-order model to establish a baseline. Following this comparison, possibilities for improvement will be suggested, and a preliminary model proposed.

## II. Methodology

This section describes the methodology followed to investigate the use of NNs for premixed MILD combustion mode identification. Two different approaches for predicting the combustion mode are followed, namely the zeroth order model and the NN. The results from both predictions are compared to the filtered binary field to determine the validity of the NN.

The methods section is structured as follows. Firstly, a description of the DNS data is given in Section II.A. Next, the data filtering and boundary exclusion processes are addressed in Section II.B. Lastly, Section II.C and Section II.D provide an explanation of the two methods.

### A. Description of DNS Data

For the present study, three cases of DNS data are used, namely A1, A2 and B1, obtained from Minamoto et al. [17]. These cases are based on a methane-air mixture, where the exhaust gases and the reactants are partially premixed in an EGR configuration. This data employed 16 species and 36 elementary reactions in a SENGA2 numerical code simulation [18].

The cases A1 and A2 entail equal EGR dilution levels, while the B1 case is characterised by higher dilution [9]. Additionally, the level of turbulence of A1 and B1 is equivalent, but is lower for A2 [9]. Case B1 is the most demanding of the three datasets due to its high dilution and high turbulence, thus, the analysis in this paper revolves around the B1 case.

The computational domain has dimensions of  $L_x \times L_y \times L_z = 10.0^3$  mm<sup>3</sup>. These domains are discretised using  $384^3$  mesh points for the B1 case, and  $512^3$  mesh points for the A1 and A2 cases, as reported by Jigjid et al. [9].

<sup>0</sup>This article was written within the course Test, Analysis and Simulation (AE2224-I) of the Aerospace Faculty of Delft University of Technology. Total word count: 4350. When writing this text, we used generative AI for inspiration and/or to structure the text and to improve the grammar, style and spelling of the text.

## B. Data preprocessing

The DNS data must be processed before its implementation. This includes filtering DNS input fields to emulate LES data using a Gaussian filter, computing the filter width sizes, as well as the exclusion of boundaries to not take into account the inlet flow conditions.

### 1. Data filtering

For correct implementation of the zeroth-order model and the NN, discussed in Sections II.C and II.D respectively, a Gaussian filter is applied to the DNS input fields to emulate LES data. This is achieved by removing SGS phenomena [16]. The implementation of the Gaussian filter is prescribed in Equation (1) [19].

$$G(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where  $\sigma$  represents the standard deviation of the function and governs the degree of smoothing, while  $x$  and  $y$  represent the grid point coordinates. A Gaussian filter is used, as it leads to gentler smoothing than a similarly sized mean filter. For higher values of  $\sigma$ , data at further distance influences the filter output, thus the field will appear more blurred.

The standard deviation is given by Equation (2), with  $k$  being the kernel size for the Gaussian filter.

$$\sigma = \sqrt{\frac{k^2}{12}} \quad (2)$$

In addition to Gaussian filtering, Favre filtering is used in Section II.D. This type of filtering serves as a method similar to Gaussian filtering, with the inclusion of local density-weighing [20].

### 2. Filter Width Size

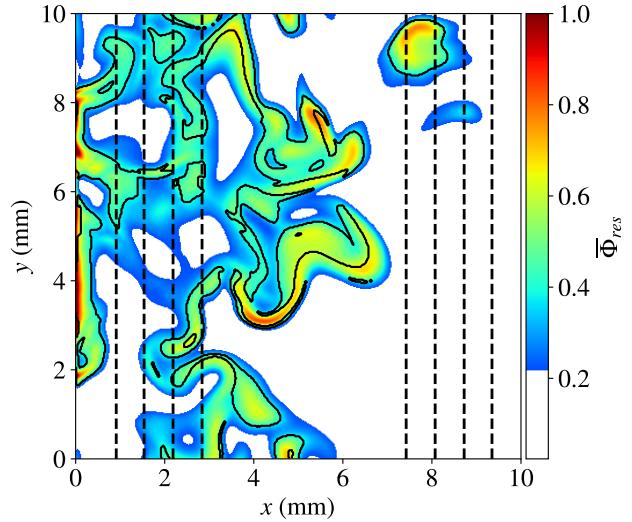
A linear relation between the laminar flame thickness and the filter width size can be formulated. The mesh width,  $\delta_x$ , is given by the simulation domain width, 10 mm, divided by the number of grid spacings [9]. Since this is the smallest division of the grid, it was selected as the basis for the filter width,  $k$ , defined by  $c\delta_x$ , where  $c \geq 1$  is the ratio of filter width to the filter width basis. The flame thickness,  $\delta_{th}$ , is constant for a particular case, with a value of 1.29 mm for B1 and 1.00 mm for A1/A2 [21]. Equating  $\delta_x$  and  $\delta_{th}$  for case B1, the linear approximation expressed by Equation (3) is retrieved:

$$50\delta_x \approx 1\delta_{th} \quad (3)$$

It is conventional for the basis of filter size to be defined in terms of flame thickness. Therefore, assume an arbitrary  $\Delta = 25k = 25c\delta_x$ . Values between 0.5 and 2.0  $\Delta/\delta_{th}$  for filtering are used, serving as input for Equation (2).

### 3. Exclusion of boundaries

Since the left and right boundary conditions have been defined artificially, the analysis of the combustion behaviour near those regions offers limited insight, and will therefore be excluded. A fixed number of cells, i.e., a base, is always excluded from the left side of the flow field as the artificial boundaries have a greater effect at the inlet. An additional cell layer is also excluded from both edges and depends on the chosen filter size, where larger kernel sizes result in a larger exclusion. This is done to minimise the representation of the inlet and outlet values in non-boundary regions, such that more points are excluded from the boundaries for larger kernels. The resulting variation in boundary exclusion for different filter sizes is visualised in Figure 1.



**Figure 1:** The instantaneous reaction rate field  $\omega_{cT}^*$  on the spatial flow field for B1. The solid contour line shows the binary combustion mode  $\Phi_{res} = 1$  region. The dashed lines show the sampling volume used for the data analyses for different filter sizes ( $\Delta/\delta_{th} = 0.5, 0.75, \dots, 2.0$ ).

## C. Definition of local combustion mode

The data used consisted of two 3D DNS data fields -  $Q/c_p$  and  $\nabla T$ , where  $Q$  is the heat release rate,  $c_p$  is the specific heat capacity at constant pressure, and  $\nabla T$  is the temperature gradient. Several steps were followed to process the input 3D field and output the binary scalar data field.

The process starts with computing the reaction rate field  $\omega_{cT}$ , obtained using Equation (4),

$$\omega_{cT} = \frac{Q}{c_p(T_b - T_u)} \quad (4)$$

With  $Q/c_p$  being a 2D DNS data field,  $T_u$  being reactant temperature,  $T_u = 1500 K$  for all cases, and  $T_b$  being burnt gas temperature,  $T_b = 1623.47 K$  for B1 and  $T_b = 1691.88 K$  for A1 and A2.  $T_u$  and  $T_b$  were based on values used in the research conducted by Minamoto

et al. [17]. Next, the  $|\nabla c_T|$  field was constructed using Equation (5).

$$|\nabla c_T| = \frac{|\nabla T|}{T_b - T_u} \quad (5)$$

With  $|\nabla T|$  being a DNS data field describing the temperature gradient.

Generalising further, the fields were normalised using their maximum values. For example, the normalised  $\omega_{cT}$  field,  $\omega_{cT}^*$ , was obtained using Equation (6).

$$\omega_{cT}^* = \frac{\omega_{cT}}{\max(\omega_{cT})} \quad (6)$$

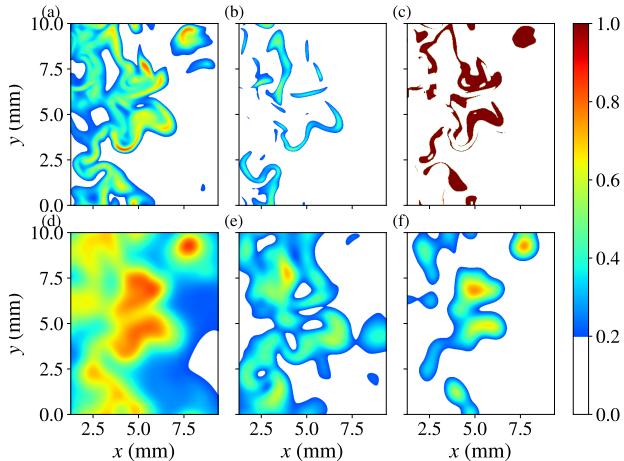
The resolved binary scalar field,  $\Phi_{res}$ , is obtained directly from the normalised DNS data fields using Equation (7) [9].

$$\Phi(\omega, \psi) = \begin{cases} 1 & \text{if } \omega_{cT}^* > 0.4 \text{ and } |\nabla c_T|^* < 0.2 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

This definition is used to obtain predicted values of  $\Phi$ . The spatial flow fields discussed in this section are visualised by Figure 2.

Lastly, the unfiltered binary field,  $\Phi_{res}$ , computed using Equation 7, is filtered as described in Section II.B.1, to obtain a filtered continuous field,  $\bar{\Phi}_{res}$ .

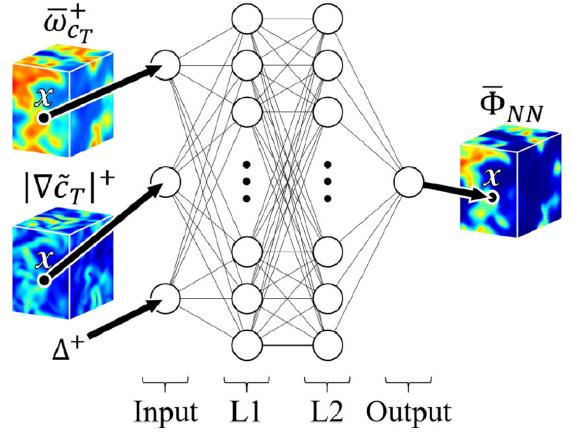
Additionally, a zeroth-order model is introduced to compare with the NN method, and consequently, quantify its validity. Such a model is implemented by applying a Gaussian filter on the DNS data before using the definition of the combustion mode, Equation (7), and then filtering the binary scalar data field to obtain  $\bar{\Phi}_{0th}$ .



**Figure 2:** Spatial flow fields for (a)  $\omega_{cT}^*$ , (b)  $|\nabla c_T|^*$  and (c)  $\Phi_{res}$  for the premixed case B1 at  $\Delta = \delta_{th}$ . The corresponding filtered fields, (d)  $\bar{\omega}_{cT}^*$ , (e)  $|\nabla \tilde{c}_T|^*$  and (f)  $\bar{\Phi}_{res}$  are also shown.

#### D. Present neural network

As an alternative to the method explained in Section II.C, where  $\Phi_{res}$  and  $\Phi_{0th}$  were directly obtained from the DNS data, an NN was trained to be used on LES data, emulated by processed DNS  $\bar{\omega}_{cT}^+$  and  $|\nabla \tilde{c}_T|$  fields. The NN contains a two-layer structure as shown in Figure 3.



**Figure 3:** Schematic of NN structure with point-wise input and output quantities at the position in the combustion domain [9].

Initially, the data obtained is processed to match the input parameters of the NN. This involved non-dimensionalising the fields, as the model inputs were selected such that the output is valid irrespective of the variance of combustion parameters, such as laminar flame thickness, reactant gas density and laminar flame speed. Equation (8) describes the non-dimensionalisation of the reaction rate:

$$\bar{\omega}_{cT}^+ = \frac{\bar{\omega}_{cT} \delta_{th}}{\rho_r S_L} \quad (8)$$

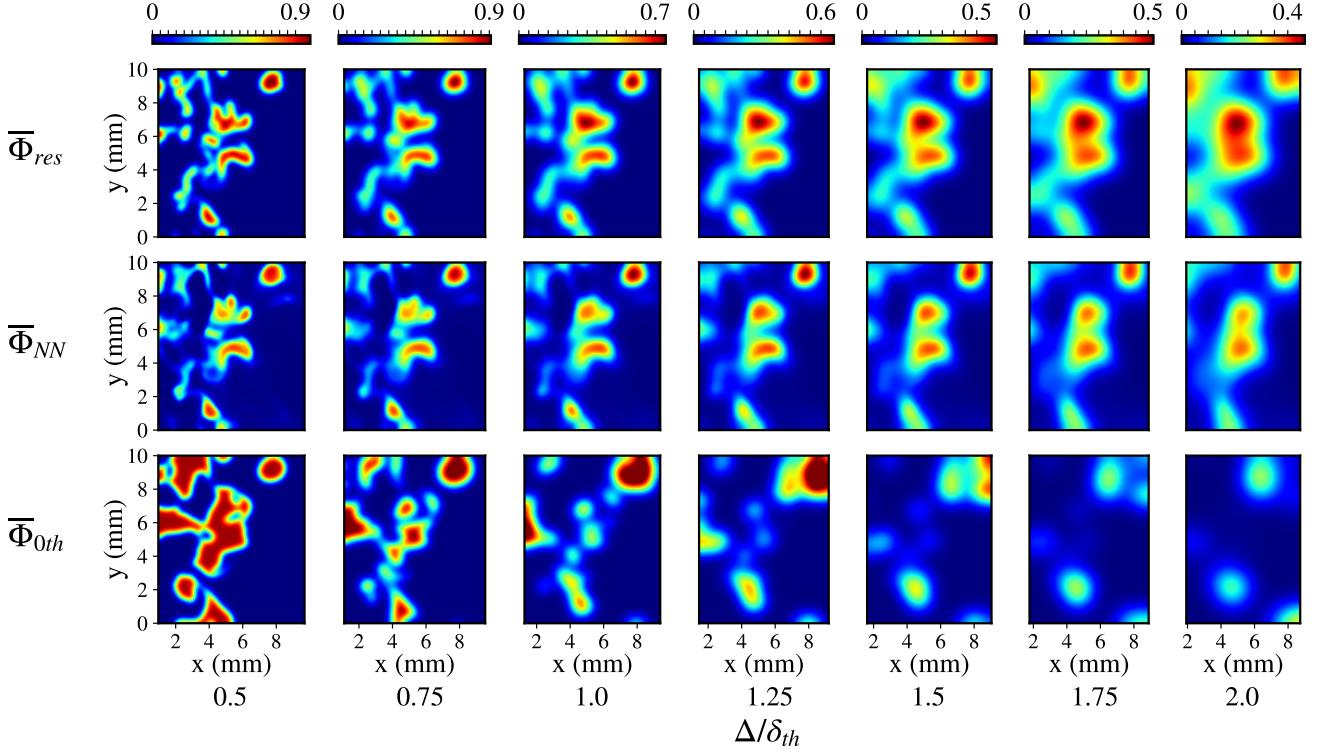
where  $\rho_r$  is reactant gas density,  $S_L$  laminar flame speed and  $\delta_{th}$  laminar flame thickness.

Similarly, the non-dimensionalised and Favre-filtered form of the normalised temperature gradient,  $|\nabla \tilde{c}_T|$  is described by Equation (9):

$$|\nabla \tilde{c}_T|^+ = \frac{|\nabla \tilde{T}|}{T_b - T_u} \delta_{th} \quad (9)$$

where the  $|\nabla \tilde{T}|$  refers to the Favre-filtered temperature gradient field.

The non-dimensionalised input fields,  $\bar{\omega}_{cT}^+$  and  $|\nabla \tilde{c}_T|^+$ , were inserted into the NN alongside the prescribed set of filter sizes to obtain the combustion mode  $\bar{\Phi}_{NN}$ , as shown in Figure 3. Section III dives deeper into the obtained results from the NN and how they match up to the methods described previously.



**Figure 4:** Visualisation of resulting filtered flow fields for  $\bar{\Phi}_{res}$ ,  $\bar{\Phi}_{NN}$  and  $\bar{\Phi}_{0th}$ . Spatial fields with increasing excluded boundaries, including filtered variation of  $\bar{\Phi}$  within the flow field (B1), for increased filter ( $\Delta/\delta_{th} = 0.5, 0.75, \dots, 2.0$ ).

### III. Results and Discussion

The validity of the NN is evaluated for the identification of combustion modes within MILD combustion following the methodology described by Section II. This is accomplished by comparing its results with those from traditional methods in Section III.A. Consequently, observations and hypotheses regarding its limitations are analysed in Section III.B, after which Section III.C proposes potential solutions.

#### A. Model results and performance

In this section, the model results are presented, obtained using the methods explained in Section II. From the model predictions, the relation between input and output features is examined, followed by quantification of the model performance.

##### 1. Model predictions

The results compared in Figure 4 are the filtered spatial fields of combustion mode  $\bar{\Phi}$  for the zeroth-order method,  $\bar{\Phi}_{0th}$ , the resolved method,  $\bar{\Phi}_{res}$ , and the NN method,  $\bar{\Phi}_{NN}$ .

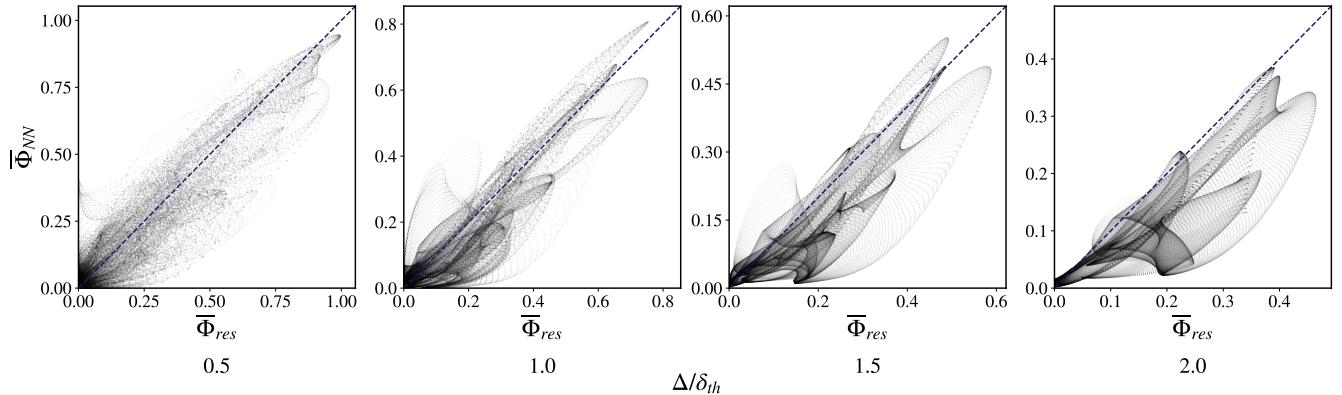
As indicated by the colour bar, regions coloured closer to crimson have higher  $\bar{\Phi}$  values, therefore dominated by the ignition/interaction combustion mode, while regions

coloured closer to indigo have lower  $\bar{\Phi}$  values, consequently being dominated by propagation or non-reacting regions. Furthermore, the domain of the  $\bar{\Phi}$  field gets narrower as the filter size increases, due to the variation in boundary exclusion on the left and right sides. This is because smaller filter sizes are affected by the inlet/outlet effects only at closer proximities, requiring smaller exclusions as explained in Section II.B.3. Moreover, the maximum  $\bar{\Phi}$  values decrease as filter size increases, due to the greater smoothening effect.

##### 2. Relation between model inputs and outputs

Based on qualitative analysis through visual comparison in Figure 4, it is evident that the NN outperforms the zeroth-order model across all filter sizes as it is closer to the  $\bar{\Phi}_{res}$  output. Additionally, insights can be gained from qualitative analyses. One notable comparison is between  $\bar{\Phi}_{NN}$  and  $\bar{\Phi}_{res}$  through point-value analysis, as illustrated in Figure 5. If the point-value analysis yields a linear trend of slope one, it indicates a direct positive correlation between  $\bar{\Phi}_{NN}$  and  $\bar{\Phi}_{res}$ .

The variations visible in Figure 5 can be attributed to the filtering done on the resulting  $\bar{\Phi}_{res}$ , as the structural patterns in the scatterplot appear more prevalent with increasing filter size. This clearly shows the pattern recognition behaviour of the NN, shown to intensify with increasing filter size.

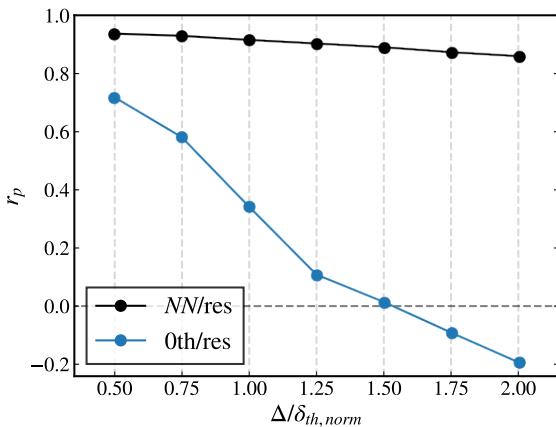


**Figure 5:** Gridpoint scatterplot depicting the filtered variation of combustion mode  $\bar{\Phi}$  within the analysed flow field for the B1 case. Correlation between DNS and NN is shown as a function of filter size  $\Delta/\delta_{th}$ .

### 3. Quantifying the model performance

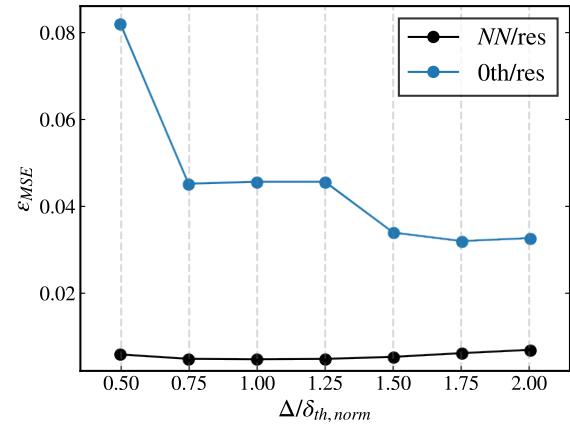
To quantify the error between the resolved data and the NN data, the Mean-Squared Error,  $\varepsilon_{MSE}$ , is calculated. Since it defines the absolute error between each point, it is widely used in various cases. However, datasets can have large absolute errors while exhibiting the same patterns, so the Pearson correlation coefficient,  $r_p$ , is also calculated between the data sets. Ideally, all values predicted by the NN are equivalent to the points modelled in the resolved model, giving a linear correlation. From these metrics, the NN shows significant potential over the zeroth-order model, as it outperforms on both Pearson coefficient and MSE at every filter size, shown in Figures 6 and 7, respectively.

Figure 6 shows that the correlation between the NN and resolved data is consistently higher than that of the zeroth-order method. Interestingly, the Pearson correlation coefficient  $r_P$  shows a slight decrease with filter size, while the MSE initially decreases and then increases. This corresponds with scientific literature [9, 16].



**Figure 6:** Pearson correlation coefficient  $r_P$  computed for  $NN/res$  (black), and  $0th/res$  (blue) obtained for case B1. Filter sizes  $\Delta/\delta_{th,norm}$  ranging between 0.5 and 2.0.

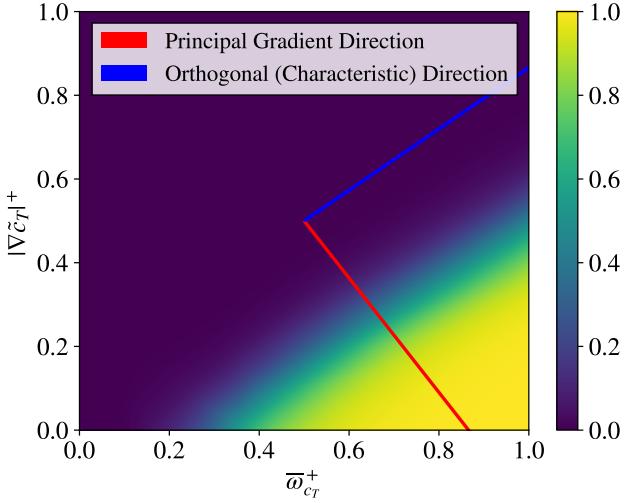
As can be seen in Figure 7, the MSE of the NN decreases for increasing filter size until  $\Delta/\delta_{th} = 1.25$ . Notably, the error increases as filter sizes reach values above 1.5. Consequently, even proportionally equivalent variations yield larger errors. Moreover, this observation shows the dependency of NN predictions on filter size optimisation. Specifically, the NN aspires to optimise predictions across all values, with an emphasis on the extremes. However, the model shows potential for improvement. Section III.B analyses the observed limitations and their causes.



**Figure 7:** Mean Squared Error  $\varepsilon_{MSE}$  computed for  $NN/res$  (black), and  $0th/res$  (blue) for the B1 case. Filter sizes  $\Delta/\delta_{th,norm}$  ranging between 0.5 and 2.0.

### B. Limitations of literature neural network

In Section III.A, it was concluded that the NN model presented significant potential for identifying combustion modes, with the analysis demonstrating the vast improvements over the zeroth-order model. Nevertheless, the model has significant limitations. In this section, the underlying predictor function, the characteristics of the



**Figure 8:** NN response  $\bar{\Phi}_{NN}$  over  $\bar{\omega}_{ct}^+$  and  $|\nabla \tilde{c}_T|^+$  for  $\Delta/\delta_{th}$ . The principal gradient direction (red) at  $-53.8^\circ$  and orthogonal direction (blue) show NN trends. PCA explains 98.95% of the variance, underscoring the significance of these directions.

training data, and the model architecture are evaluated.

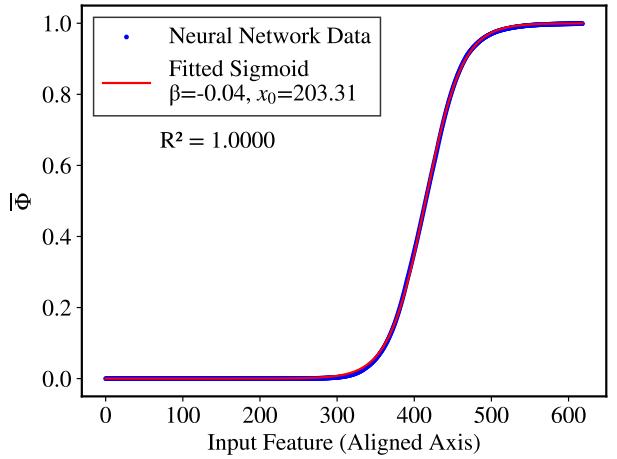
### 1. Underlying predictor function

This section explores the predictions of the NN to a range of input values for  $\bar{\omega}_{ct}^+$  and  $|\nabla \tilde{c}_T|^+$  to interpret the existence and/or properties of an underlying prediction function. The motivation for this analysis stems from initial observations suggesting that the network may have difficulty accurately capturing intermediate values, as seen in Figure 5, potentially limiting its effectiveness.

In order to evaluate the response to any combination of inputs, Principal Component Analysis (PCA) is used. PCA is a statistical technique used to reduce the dimensionality of a dataset while retaining the information of the principal components, i.e., linear combinations of the original variables capturing the maximum data variance [19]. In the context of analysing the output of the NN, PCA helps to reveal the dominant relationship between the input variables  $\bar{\omega}_{ct}^+$  and  $|\nabla \tilde{c}_T|^+$ .

As shown in Figure 8, the application of PCA to the NN's output reveals that about 99% of the data variance is explained by the principal components. This percentage indicates that the NN's behaviour is largely determined by a few dominant patterns, visualised as the principal gradient direction and the orthogonal direction in Figure 8. While these dominant patterns provide insights into the NN's basic operation, they also suggest that the NN may be oversimplifying the dynamics of the combustion process.

To further investigate this simplification, a sigmoidal function is fitted to the NN's output along the principal gradient direction, shown in Figure 9. The perfect fit ( $R^2 = 1.00$ ) of this sigmoidal curve demonstrates that the NN's behaviour along this direction can be



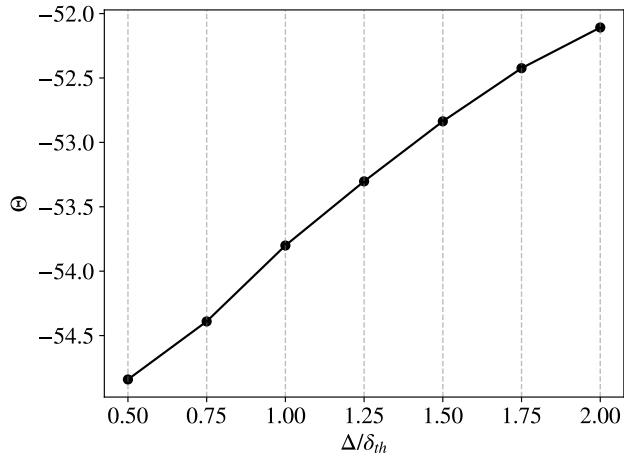
**Figure 9:** Sigmoid fit to neural network output along the principal axis for  $\Delta/\delta_{th}$ , modelled by Equation (10). This plot illustrates the model's predictable behaviour with an excellent fit ( $R^2 = 1.00$ ), highlighting  $\beta$  and  $x_0$  as the optimised parameters of the curve.

accurately described by a simple mathematical function, denoted by Equation (10).

$$\bar{\Phi} = \frac{1}{1 + e^{-\beta(x-x_0)}} \quad (10)$$

However, the combustion process is likely to involve more intricate interactions and local variations that cannot be fully captured by a single sigmoidal curve. The use of a sigmoid activation function in the output layer of the NN may contribute to this oversimplification.

Figure 10 illustrates the linear increase in the angle of the principal gradient direction as the filter size increases. This behaviour indicates that the NN's learning process is disproportionately influenced by the characteristics of the filtered data.



**Figure 10:** Variation of the principal gradient direction angle  $\Theta$  with respect to the horizontal, counterclockwise positive, as a function of  $\Delta/\delta_{th}$ . The angle shifts from  $-54.8^\circ$  to  $-52.1^\circ$  across the range, representing a small but consistent percentage change with increasing  $\Delta/\delta_{th}$ .

As larger filters incorporate information from broader data regions, they modify the underlying gradient structures, potentially altering the NN's perception of the principal directions.

The consistent change in the principal direction angle with increasing filter size indicates that the network may be overly adapted to data smoothing characteristics, rather than capturing the fundamental and invariant features of the combustion process, such as the temperature distribution, chemical reaction rates, and the formation of intermediate species.

These findings underscore the importance of refining the model training strategies to ensure more robust and accurate predictions across a range of operational conditions.

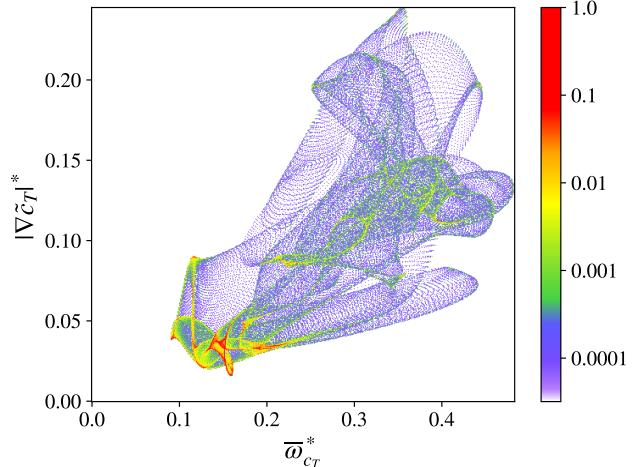
## 2. Training data characteristics

The simplicity of the NN's response can be attributed to multiple factors, including the spatial filtering and imbalance of the training dataset.

The spatial filtering of the binary  $\Phi_{res}$  field using a Gaussian filter to obtain a continuous  $\bar{\Phi}_{res}$  field in the  $x - y$  plane causes significant diffusion between the  $\bar{\Phi}_{res}$  values in the  $|\nabla \tilde{c}_T|^+ - \bar{\omega}_{ct}^+$  plane. This filtering process does not consider the neighboring  $|\nabla \tilde{c}_T|^+$  and  $\bar{\omega}_{ct}^+$  values, leading to situations where points with approximately the same  $|\nabla \tilde{c}_T|^+$  and  $\bar{\omega}_{ct}^+$  values exhibit large variations in  $\bar{\Phi}_{res}$ , as shown in Figure 12. Consequently, the NN, which is a deterministic point-to-point method [22], encounters difficulties in providing a one-to-one mapping between the input,  $|\nabla \tilde{c}_T|^+$  and  $\bar{\omega}_{ct}^+$ , and output,  $\bar{\Phi}_{res}$  in certain regions. This limitation hinders the NN's ability to capture the underlying combustion dynamics accurately.

This inconsistent mapping, resulting from the diffusive effects of the Gaussian filtering, can be interpreted as the target function,  $\bar{\Phi}_{res}$ , becoming non-deterministic. Although the Universal Approximation Theorem is typically stated for continuous functions, it can also be applied to discrete datasets, as continuous functions can be discretized. However, this non-deterministic target function violates the continuity requirement of the theorem, as stated by Leshno et al. [23], which asserts that "a standard multilayer feedforward network with a locally bounded piecewise continuous activation function can approximate any continuous function to any degree of accuracy if and only if the network's activation function is not a polynomial." Consequently, the NN's performance is adversely affected, despite its architecture satisfying the theorem's conditions.

An additional observation from Figure 5 is that larger errors generally occur at higher  $\bar{\Phi}_{res}$  values. This can be attributed to the imbalance of the dataset, i.e., the lack of training data for high  $\bar{\omega}_{ct}^*$  and low  $|\nabla \tilde{c}_T|^*$  ( $\Phi_{res} = 1$ ) compared to low  $\bar{\omega}_{ct}^*$  or high  $|\nabla \tilde{c}_T|^*$  ( $\Phi_{res} = 0$ ) [24], as shown in Figure 11.



**Figure 11:** Normalised frequency plot of occurrences of  $\bar{\omega}_{ct}^*$  and  $|\nabla \tilde{c}_T|^*$  combinations within the spatial field for case B1. Filter size depicted is  $1.25\Delta/\delta_{th}$ , yet represents phenomena across all filter sizes. Note the logarithmic scale of the color bar.

The imbalanced dataset has two main consequences. Firstly, as mentioned in Section III.A, the filtering process has a greater effect on these regions, resulting in a more pronounced non-deterministic effect. Secondly, the limited amount of data in these regions hinders the ability of the NN to adapt and accurately predict the combustion mode for those conditions [24].

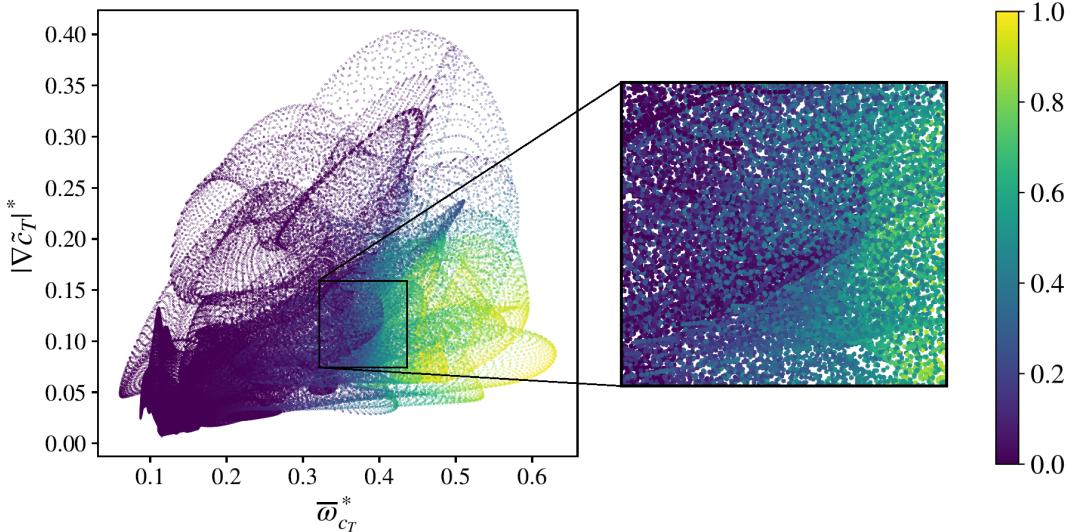
## 3. Model architecture

The current NN architecture, consisting of two hidden layers with ReLU activation functions and a sigmoidal output function, meets the requirements of the Universal Approximation Theorem, as discussed in Section III.B.2. However, there are certain aspects of the architecture that could be improved to enhance the NN's performance and generalisation ability.

One such aspect is the depth of the network. While the current architecture with two hidden layers has shown improved performance compared to the zeroth-order model, evident from the results presented in Figures 6 and 7, deeper architectures are often more effective in capturing complex hierarchical features and abstractions [25, 26].

Another limitation lies in the choice of activation functions. The use of ReLU activation functions in the hidden layers, while computationally efficient and effective in mitigating the vanishing gradient problem [27], may not be the most suitable choice for this application. ReLU functions are piecewise linear and have zero output for negative inputs, which can lead to reduced representational capacity for more complex functions in shallow networks [28].

Furthermore, the architecture does not incorporate regularisation techniques, such as L1/L2 regularisation or dropout [29], which can help prevent overfitting and improve generalisation ability. Given the non-deterministic nature of the inputs caused by the filtering process,



**Figure 12:**  $\bar{\Phi}_{res}$  for inputs  $\bar{\omega}_{cT}^*$  and  $|\nabla \tilde{c}_T|^*$  for B1 data, with filter size of  $0.5 \Delta/\delta_{th}$ . Note the logarithmic scale of the colour bar.

the impact of regularization techniques on the NN’s performance is uncertain and should be investigated empirically.

In summary, the primary limitation in the NN’s performance stems from the non-deterministic nature of the training data, which violates the continuity assumption required for the Universal Approximation Theorem. While addressing these discontinuities is crucial for improving the NN’s approximation capabilities, potential areas for improvement in the NN architecture include increasing network depth, exploring alternative activation functions, and incorporating regularisation techniques. By combining data preprocessing techniques to mitigate the effects of non-deterministic characteristics with architectural optimisations, the NN’s ability to capture the dynamics of the combustion process and generalise to unseen data can be significantly enhanced.

### C. Potential Improvements

Having identified the limitations of the current NN architecture, this section proposes potential improvements. As mentioned in Section III.B the main limitations of the NN are the activation functions, the training data, and the depth of the network.

#### 1. Training data

To address the non-deterministic characteristics of the training data, it is suggested to discretise the  $|\nabla \tilde{c}_T|^+$  and  $\bar{\omega}_{cT}^+$  fields onto a coarser basis and compute the average value at each discretised point. By training the NN on these expected values, which are deterministic, the issues related to the non-deterministic characteristics of the training data can be mitigated, allowing the NN to better capture the underlying combustion dynamics. Moreover, discretising into a coarser basis

can also reduce the effects of an imbalanced dataset, as all regions would be represented equally, allowing the NN to effectively adjust to general trends.

#### 2. Activation functions

As discussed in Section III.B.3, the use of ReLU activation functions in the hidden layers may not be the most suitable choice for capturing the highly non-linear characteristics of the combustion process [21]. Alternative activation functions, such as the hyperbolic tangent (tanh) or sigmoid function, could be more appropriate for introducing the necessary non-linearities in the model and improving its representational capacity [30].

#### 3. Depth of the network

Increasing the number of hidden layers in the NN architecture could improve its ability to capture the complex relationships and non-linearities present in the combustion process. A deeper architecture, with at least three hidden layers, could ensure better adjustments to non-linear components, enhancing the NN’s capacity to represent the intricate dynamics of the combustion process. Furthermore, a greater dataset is recommended to account for the imbalance of the dataset and allow for greater generalisation of the NN.

A preliminary model that incorporates the aforementioned improvements was developed and tested. The results, shown in Figures 14 and 15, demonstrate better performance than the original model, with higher Pearson coefficients and lower MSE values across all filter sizes. This suggests that addressing the identified limitations can significantly enhance the NN’s performance for combustion mode identification in MILD combustion. A detailed explanation of the improved NN architecture and its performance is presented in Appendix A.

## IV. Conclusion

This study investigated and improved the usage of an existing Neural Network (NN) by Jigjid et al. [9] for the identification of combustion modes under Moderate or Intense Low-oxygen Dilution (MILD) combustion.

Quantitative analysis revealed that the NN outperforms the zeroth-order model across all filter sizes in both Pearson correlation coefficient and MSE. While the NN predictions seemed to correlate with DNS outcomes, PCA revealed that the behaviour of the NN was largely governed by a few dominant patterns, oversimplifying the complex combustion dynamics. This is assumed to be mainly due to the incompatibility between the non-deterministic characteristics of the training data and the deterministic behaviour of the NN, further exacerbated by choice of activation functions and the depth of the network.

An improved NN was proposed and implemented by discretising the data into a coarser basis, using the expected values per discretised location, and incorporating a deeper architecture with three hidden layers and sigmoid activation functions. The revised NN yielded significantly better results compared to the original architecture. It is recommended to investigate the physical applicability of the improved NN to other datasets to assess its ability to generalise.

This study highlights the importance of larger datasets for successfully training NNs and achieving better results. Addressing the limitations related to the model architecture, such as the shallow depth and choice of activation functions, and the characteristics of the training data, including the non-deterministic behaviour, and an imbalanced dataset, can potentially lead to significant improvements in the NN's performance and reliability for MILD combustion mode identification. Future research into the use of CNNs for these applications is also recommended.

## Acknowledgments

This study was supported by the AE2224-I Test, Analysis & Simulation course of the Delft University of Technology. The authors would like to thank K. Jigjid and Dr. N.A.K. Doan for providing technical advice throughout the process of composing this paper.

## References

- [1] Noor, M., Wandel, A. P., and Yusaf, T., "A review of mild combustion and open furnace design consideration," *International Journal of Automotive and Mechanical Engineering (IJAME)*, Vol. 6, 2012, pp. 730–754. doi: 10.15282/ijame.6.2012.6.0060.
- [2] Cavaliere, A., and de Joannon, M., "Mild Combustion," *Progress in Energy and Combustion Science*, Vol. 30, No. 4, 2004, pp. 329–366. doi:10.1016/J.PECS.2004.02.003.
- [3] Sorrentino, G., Sabia, P., Bozza, P., Ragucci, R., and de Joannon, M., "Impact of external operating parameters on the performance of a cyclonic burner with high level of internal recirculation under MILD combustion conditions," *Energy*, Vol. 137, 2017, pp. 1167–1174. doi:10.1016/J.ENERGY.2017.05.135.
- [4] Sorrentino, G., Cavaliere, A., Sabia, P., Ragucci, R., and de Joannon, M., "Diffusion Ignition Processes in MILD Combustion: A Mini-Review," *Frontiers in Mechanical Engineering*, Vol. 6, 2020. doi:10.3389/fmech.2020.00010.
- [5] Perpignan, A. A., Sampat, R., and Gangoli Rao, A., "Modeling Pollutant Emissions of Flameless Combustion With a Joint CFD and Chemical Reactor Network Approach," *Frontiers in Mechanical Engineering*, Vol. 5, 2019. doi:10.3389/fmech.2019.00063.
- [6] Kruse, S., Kerschgens, B., Berger, L., Varea, E., and Pitsch, H., "Experimental and numerical study of MILD combustion for gas turbine applications," *Applied Energy*, Vol. 148, 2015, pp. 456–465. doi:10.1016/j.apenergy.2015.03.054.
- [7] Wang, F., Li, P., Mi, J., and Shu, Z., "A novel method to improve stability of MILD combustion in a highly heat-extracted furnace," *Fuel*, Vol. 292, 2021, p. 120315. doi: 10.1016/J.FUEL.2021.120315.
- [8] Doan, N. A. K., and Swaminathan, N., "Autoignition and flame propagation in non-premixed MILD combustion," *Combustion and Flame*, Vol. 201, 2018, pp. 234–243. doi: 10.1016/j.combustflame.2018.12.025.
- [9] Jigjid, K., Tamaoki, C., Minamoto, Y., Nakazawa, R., Inoue, N., and Tanahashi, M., "Data driven analysis and prediction of MILD combustion mode," *Combustion and Flame*, Vol. 223, 2021, pp. 474–485. doi:10.1016/J.COMBUSTFLAME.2020.10.025.
- [10] Wünning, J. A., and Wünning, J. G., "Flameless oxidation to reduce thermal no-formation," *Progress in Energy and Combustion Science*, Vol. 23, No. 1, 1997, pp. 81–94. doi: 10.1016/S0360-1285(97)00006-3.
- [11] Minamoto, Y., and Swaminathan, N., "Subgrid scale modelling for MILD combustion," *Proceedings of the Combustion Institute*, Vol. 35, No. 3, 2015, pp. 3529–3536. doi:10.1016/J.PROCI.2014.07.025.
- [12] Egolfopoulos, F. N., Hansen, N., Ju, Y., Kohse-Höinghaus, K., Law, C. K., and Qi, F., "Advances and challenges in laminar flame experiments and implications for combustion chemistry," *Progress in Energy and Combustion Science*, Vol. 43, 2014, pp. 36–67. doi:10.1016/j.pecs.2014.04.004.
- [13] Kayne, A., and Agarwal, R. K., "Computational fluid dynamics modeling of mixed convection flows in buildings enclosures," *International Journal of Energy and Environment*, Vol. 4, No. 6, 2013, pp. 2076–2909. doi:10.1115/ES2013-18026.
- [14] Li, Z., Tomasch, S., Chen, Z. X., Parente, A., Ertesvåg, I. S., and Swaminathan, N., "Study of MILD combustion using LES and advanced analysis tools," *Proceedings of the Combustion Institute*, Vol. 38, No. 4, 2021, pp. 5423–5432. doi:10.1016/J.PROCI.2020.06.298.
- [15] Doan, N. A. K., Swaminathan, N., and Minamoto, Y., "DNS of MILD combustion with mixture fraction variations," *Combustion and Flame*, Vol. 189, 2018, pp. 173–189. doi: 10.1016/j.combustflame.2017.10.030.

- [16] Jigjid, K., Minamoto, Y., Doan, N. A. K., and Tanahashi, M., “SGS Reaction rate modelling for MILD combustion based on machine-learning combustion mode classification: Development and a priori study,” *Proceedings of the Combustion Institute*, Vol. 39, No. 4, 2023, pp. 4489–4499. doi:10.1016/J.PROCI.2022.07.020.
- [17] Minamoto, Y., Swaminathan, N., Cant, R. S., and Leung, T., “Morphology of reaction zones in MILD and premixed combustion,” *Combustion and Flame*, Vol. 161, No. 11, 2014, pp. 2801–2814. doi:10.1016/j.combustflame.2014.04.018.
- [18] Cant, S., and University of Cambridge. Engineering Department, *SENGA2 User Guide*, cued/a-thermo/tr ed., University of Cambridge, Department of Engineering, 2012.
- [19] Pedregosa, F., Michel, V., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Varoquaux, G., Gramfort, A., Thirion, B., Dubourg, V., Passos, A., Brucher, M., and Duchesnay, E., “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2825–2830. URL <http://jmlr.org/papers/v12/pedregosa11a.html>.
- [20] Lesieur, M., and Comte, P., “Favre filtering and macro-temperature in large-eddy simulations of compressible turbulence,” *Comptes Rendus de l’Académie des Sciences - Series IIB - Mechanics*, Vol. 329, 2001, pp. 363–368. doi:10.1016/S1620-7742(01)01331-9.
- [21] Minamoto, Y., “Physical Aspects and Modelling of Turbulent MILD Combustion,” Ph.D. thesis, 2013.
- [22] Brouwer, R. K., *ESANN’2002 : 10th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 24-25-26, 2002, proceedings*, D-side, 2002.
- [23] Leshno, M., Lin, V. Y., Pinkus, A., and Schocken, S., “Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function,” Tech. rep., 1993.
- [24] Wang, S., Liu, W., Wu, J., Cao, L., Meng, Q., and Kennedy, P. J., “Training deep neural networks on imbalanced data sets,” *Proceedings of the International Joint Conference on Neural Networks*, Vol. 2016-October, Institute of Electrical and Electronics Engineers Inc., 2016, pp. 4368–4374. doi:10.1109/IJCNN.2016.7727770.
- [25] Bengio, Y., “Learning deep architectures for AI,” *Foundations and Trends in Machine Learning*, Vol. 2, No. 1, 2009, pp. 1–27. doi:10.1561/2200000006.
- [26] Lecun, Y., Bengio, Y., and Hinton, G., “Deep learning,” , 5 2015. doi:10.1038/nature14539.
- [27] Glorot, X., Bordes, A., and Bengio, Y., “Deep Sparse Rectifier Neural Networks,” Tech. rep., 2011.
- [28] Mao, T., and Zhou, D. X., “Rates of approximation by ReLU shallow neural networks,” *Journal of Complexity*, Vol. 79, No. 101784, 2023. doi:10.1016/J.JCO.2023.101784.
- [29] Srivastava, N., Hinton, G., Krizhevsky, A., and Salakhutdinov, R., “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” Tech. rep., 2014.
- [30] Haykin, S. S., and Haykin, S. S., *Neural networks and learning machines*, Prentice Hall/Pearson, 2009.

## Appendix

### A. Preliminary Model Architecture

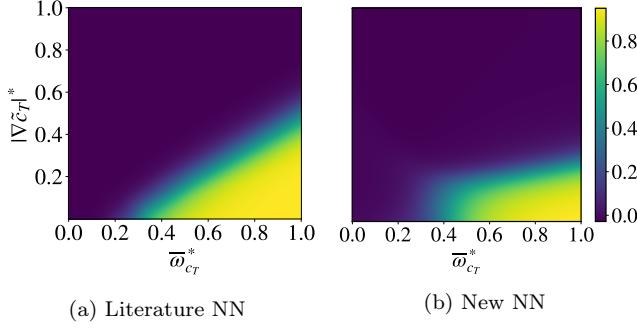
Section III presented NNs as a viable option for combustion type identification, but argued that there are substantial issues with the implementation conducted by Minamoto et al. [17]. This section aims to substantiate those arguments through the proposal of an alternative solution that addresses the issues outlined in the previous realisation. For convention purposes,  $\Phi_{NN}$  of the analysed network will be depicted by  $\Phi_{NN,lit}$ , while the new network will be addressed by  $\Phi_{NN,new}$ .

As mentioned, the main issue with the model from literature is that the target data was not deterministic, which is not valid for a deterministic NN. The new implementation will execute the solution suggested previously, of discretising the data and taking the expected value per discretised location. In the preliminary NN, multiple discretisation meshes were tested, and approximately 65 was found to be optimal for the training data. This is attributed to be the main factor when fixing issues related to the overly-simplified symbolic approximation and the under-prediction for higher  $\bar{\Phi}_{res}$  values.

Before introducing the solution, it is essential to address present considerations regarding architectural choices in existing literature. Using two hidden layers with ReLU activation functions may not suit the present needs. While ReLU is favoured over sigmoid functions, its efficacy in shallow networks is questionable due to weaker non-linear properties. This simplicity may lead to a sigmoidal shape due to a sigmoid activation in the output layer. Additionally, the impact of batch processing is unclear. While it may mitigate non-deterministic target issues, limited training data could destabilise activation changes, limiting predictive abilities in specific directions.

The preliminary NN architecture featured three layers, ensuring that every input combination incorporates a non-linear component. Sigmoid activation functions were employed to better address non-linearities within the shallow network. The NN was trained using A1 and A2 cases, with three samples each. Additionally, the network operated without batching, leveraging dense training data. Specifically, 80% of the data from each sample was used per iteration, with all data points shuffled into a single input set. The number of nodes per hidden layer set to an arbitrary value of 40, with the Adam optimisation algorithm employed for training and MSE as the loss function.

The optimal learning rate for the new NN was found to be slightly higher, but this was expected, due to the larger number of layers, and the use of sigmoidal activation functions, which have worse learning properties than ReLU activation functions.



(a) Literature NN

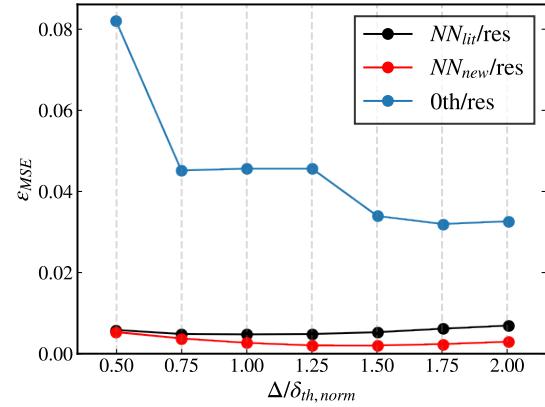
(b) New NN

**Figure 13:**  $\bar{\Phi}_{NN}$  prediction values for maximal domain of  $\bar{w}_{cr}^*$  and  $|\nabla \tilde{c}_T|^*$  for  $\Delta/\delta_{th} = 0.5$ .

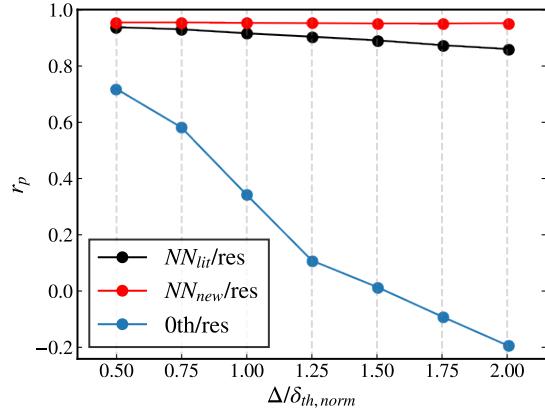
The NN was trained through 40 epochs, as further iterations diverged for the learning rate used. The results of the NN are shown in Figure 13.

Note that the new network is not particularly superior to the literature NN at the extrema, but does excel at predicting intermediate  $\bar{\Phi}$  values. It predicts more complex shapes than simple sigmoids, thus, resulting in lower  $\varepsilon_{MSE}$  and higher  $r_p$  values than the literature NN for all filter sizes, as shown in Figures 14 and 15.

It must be mentioned that multiple dropout rates were tested, but the results were significantly worse than without dropout, thus, it was not used. However, results may vary for different datasets.



**Figure 14:** Mean Square Error,  $\varepsilon_{MSE}$ , of  $\bar{\Phi}$  computed for  $NN_{lit}/res$  (black),  $NN_{new}/res$  (red), and  $0th/res$  (blue) obtained for case B1. Filter sizes  $\Delta/\delta_{th,norm}$  ranging between 0.5 and 2.0.



**Figure 15:** Pearson correlation coefficient,  $r_p$ , of  $\bar{\Phi}$  computed for  $NN_{lit}/res$  (black),  $NN_{new}/res$  (red), and  $0th/res$  (blue) obtained for case B1. Filter sizes  $\Delta/\delta_{th,norm}$  ranging between 0.5 and 2.0.

## B. Logbook

**Table 1:** Logbook of Contributions

Section	Contributor(s) <sup>1</sup>
<b>Abstract</b>	Carla
<b>Introduction</b>	Blanca, Siddarth
<b>Methodology</b>	
Description of DNS Data	Blanca, Gabriël, Job, Siddarth
Data Preprocessing	
1. Data Filtering	Blanca, Carla, Job, Siddarth
2. Filter Width Size	Blanca, Gabriël, Job, Mario, Siddarth
3. Exclusion of Boundaries	Gabriël, Job, Siddarth
Definition of Local Combustion Mode	Blanca, Carla, Job, Siddarth
Present Neural Network	Carla, Daniël, Job, Siddarth
<b>Results and Discussion</b>	
Model Results and Performance	
1. Model Predictions	Mario, Nina
2. Relation Between Model Inputs and Outputs	Gabriël, Mario, Nina
3. Quantifying the Model Performance	Gabriël, Mario, Nina and Owen
Limitations of Literature Neural Network	
1. Underlying Predictor Function	Alejandro
2. Training Data Characteristics	Alejandro, Mario
3. Model Architecture	Alejandro, Mario, Owen
Potential Improvements	
1. Training Data	Alejandro, Mario
2. Activation Functions	Alejandro, Mario, Owen
3. Depth of the Network	Alejandro, Mario, Owen
<b>Conclusion</b>	Alejandro, Owen, Siddarth
<b>Appendix A</b>	Gabriël, Mario
<b>Proofreading</b>	Everyone
<b>Programming</b>	Alejandro, Daniël, Gabriël, Mario
<b>LaTeX Formatting</b>	Alejandro, Gabriël

<sup>1</sup>Author order is alphabetical.