Accepted for publication in Applications in Energy and Combustion Science, 05/May/2023

Towards a Generalised Artificial Neural Network for Sub-grid Filtered Density Function Closure in Turbulent Combustion

Hanying Yang[a,*], Tota Kobayashi[b,a], Salvatore Iavarone[c,a], James C. Massey[a,d], Zhi X. Chen[e,f,a], Yuki Minamoto[b,1], Nedunchezhian Swaminathan[a]

[a]*Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, United Kingdom*
[b]*School of Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro, Tokyo 152-8850, Japan*
[c]*École polytechnique de Bruxelles, Aero-Thermo-Mechanics Laboratory, Université Libre de Bruxelles, Avenue F. D. Roosevelt 50, Brussels 1050, Belgium*
[d]*Robinson College, University of Cambridge, Grange Road, Cambridge CB3 9AN, United Kingdom*
[e]*State Key Laboratory of Turbulence and Complex Systems, Aeronautics and Astronautics, College of Engineering, Peking University, Beijing 100871, China*
[f]*AI for Science Institute (AISI), Beijing 100080, China*

Abstract

An artificial neural network (ANN) is trained on moderate or intense low-oxygen dilution (MILD) combustion to predict the sub-grid filtered density function (FDF) in large eddy simulation (LES). For wide usability, a new quantity ψ is calculated by scaling the filtered mixture fraction using the stoichiometric value and a logarithmic function, which is used as an input feature of the ANN. Self-testing on MILD combustion is first conducted for validation and determining the best layout of ANN. An ANN with 4 hid-

*Corresponding author
Email address: hy345@cam.ac.uk (Hanying Yang)
[1]Fixstars Amplify Corporation, 3-1-1 Shibaura, Minato, Tokyo,108-0023, Japan

den layers and the activation function of rectified linear unit (ReLU) has the highest accuracy. This ANN is subsequently tested on five different premixed combustion cases. Overall predictions of the progress variable FDF are satisfactory, and the filtered chemical source term modelled by the FDFs is in good agreement with the DNS data for each case. However, an underprediction of the filtered reaction rate occurs for the hydrogen–air flames, and the ANN accuracy is lower than that of a presumed β-FDF approach. DNS data of one premixed flame are then added to the training datasets, and newly trained ANNs are tested on remaining premixed flame cases. When the testing cases are similar to the added training dataset, improved predictions of marginal FDFs are observed, and the modelled progress variable source terms show good comparisons with the corresponding DNS data. Compared to the presumed β-FDF method, the new ANN shows an improved accuracy and serves as a viable alternative for FDF-based combustion modelling in the LES of turbulent flames.

Keywords: artificial neural network, filtered density function, MILD combustion, premixed flames

---

1. Introduction

Large eddy simulation (LES) is becoming a prevalent method for studying combustion as it offers high accuracy with an acceptable computational cost. In LES, physical processes associated with scales larger than the filter width Δ are fully resolved, while the smaller scales below the filter width, known as the sub-grid scale (SGS) range, are modelled. Such models are required to describe the sub-grid turbulent advection and diffusion [1] and the filtered reaction rates [2] in chemically reacting flows. Models for sub-grid turbulent advection and diffusion are either based on algebraic expressions or transport equations, and both are extensively studied for non-reactive [3] and reactive [4] flows. Describing reaction rates using a simple model is challenging as chemical reactions are known to be stiff, non-linear and interact strongly with turbulence [4].

Artificial neural networks (ANNs) may be a suitable alternative to model the filtered reaction rates in the LES framework. Since the 1990s, this technique has been utilised in the realm of combustion modelling, especially in predicting dynamic chemical states [5–9]. Subsequently, more complex combustion systems have been investigated. Sen *et al.* [10–12] trained their ANNs using a set of unsteady flame-turbulence-vortex-interaction cases, and successfully simulated turbulent combustion of syngas with air. Chatzopoulos *et al.* [13] used non-premixed flamelets to generate training data, Franke et al. [14] included local extinction and re-ignition conditions, and Wan *et al.* [15] included non-adiabatic flames in the training set. Tabulated flamelet-based models are also a common avenue. Generally, the thermo-chemical quantities used as control variables for the look-up table (LUT) are the ANN

input features, while the tabulated quantities are the outputs. This novel technique was initially explored by Flemming *et al.* [16], Kempf *et al.* [17] and Emami and Fard [18]. Then Ihme *et al.* [19, 20] optimized the ANN and accurately simulated the Sydney bluff-body swirl-stabilised flame. Recently, Owoyele *et al.* [21] trained ANNs by using multi-dimensional flamelet libraries and predicted the n-dodecane spray flame and methyl decanoate combustion under various conditions with high accuracy; Ranade *et al.* [22] deployed ANN in the LES of the DLR-A turbulent jet diffusion flame to supersede the 4-dimensional probability density function (PDF) table. More applications of ANN in reactive flows can be found in the previous studies [23, 24]. On the other hand, the narrow application of ANNs is an outstanding issue which is limited by the scope of the training data. A high predictive ability is observed when the conditions of the test cases are covered by the training sets, whereas the ability becomes poor otherwise. However, generating a training data repository covering these and other potential thermochemical conditions that may arise in a targeted test case is very challenging.

The presumed PDF approach in LES needs a sub-grid filtered density function (FDF), and the ANN may be leveraged to infer the FDF by using a relevant control variable as input features. This was explored by De Frahan *et al.* [25] and tested on a swirling methane–air premixed flame. They showed that ANN could be the most suitable algorithm to infer the marginal FDF of the reaction progress variable using its first two moments. Yao *et al.* [26] concluded that a properly trained ANN yielded an improved marginal FDF of mixture fraction in a turbulent spray flame compared with the conventional presumed shapes. Chen *et al.* [27] advanced the state-of-the-art by predicting

joint FDF of progress variable and mixture fraction via ANN for moderate or intense low-oxygen dilution (MILD) combustion. A superior performance of the ANN, compared to advanced presumed PDF methods, was reported in Ref. [27]. However, all of these studies are self-testing, implying that the training and testing data cover analogous turbulent and thermo-chemical conditions. This does not guarantee a good generality for ANN-based FDFs, and using completely different data sets for training and testing ANN is essential.

The present work aims to develop and train an ANN that can infer FDF in various turbulent flames. For this, The ANN is tested by using a turbulent planar, swirl, V-, twin V- and stagnation flames considering either hydrogen (H_2) or methane (CH_4) as fuels, whereas only a single MILD combustion case, with its specific thermo-chemical and turbulence conditions, is considered for training and validation. As an initial step, only premixed flames are tested in this study. Therefore, the marginal FDF of the reaction progress variable is of interest, and this work will focus on *a priori* assessment using direct numerical simulation (DNS) data, while *a posteriori* testing through LES is of future interest. The remainder of this paper is organised as follows. Section 2 describes the setup of the ANN and the DNS datasets considered for its training, validating, and testing. The FDF predicted by the ANN, along with the filtered reaction rate modelled via this FDF, will be compared with the corresponding values from DNS in Section 3, followed by further discussion on improvements in Section 4. The conclusions are summarised in Section 5.

2. ANN methodology

2.1. Training case

The dataset from DNS of MILD combustion with varying mixture fraction and internal recirculation of exhaust gases (EGR) is used to train the ANN. The presence of both premixed and non-premixed combustion is observed [28, 29] in turbulent combustion under MILD conditions, and the trained ANN is expected to give good inference for all modes of combustion. Furthermore, a broad reaction zone was observed in this MILD combustion case [28, 29], implying a wide range of subgrid-scale fluctuations in the thermo-chemical and -physical conditions of the reacting mixture. This is desirable for ANN training to ensure that a robust ANN can be constructed.

One DNS dataset, labelled 'AZ1' and described in detail in Ref. [28], is considered for this study. Both thin and thick heat release zones have been observed in this case, which correspond to a premixed mode and the interaction of the reaction zones respectively [28]. The initial thermo-chemical and -physical conditions of AZ1 are listed in Table 1. The initial velocity fluctuation u' and integral length scale Λ_0 are around 16.66 m/s and 1.42 mm respectively, yielding a turbulent Reynolds number Re_t and Taylor micro-scale Reynolds number Re_λ of about 96 and 34.73, respectively. The reaction progress variable and mixture fraction vary spatially, and the ratio of their integral length scales is $l_c/l_Z \approx 0.77$. The dilution is controlled by decreasing the maximum concentration of oxygen $X_{O_2}^{max}$ to 3.5% by volume. The corresponding stoichiometric mixture fraction value is $Z_{st} = 0.01$. The averaged mixture fraction in the DNS domain is $\langle Z \rangle = 0.008$, suggesting overall fuel-lean combustion. The averaged progress variable value is $\langle c \rangle = 0.56$, which

implies that both unburnt and burnt mixtures exist within the simulation domain. The root-mean-square values of the initial fluctuations in the mixture fraction and progress variable fields are $\sigma_Z/\langle Z\rangle = 1.05$ and $\sigma_c/\langle c\rangle = 0.46$, respectively.

A cubic computational domain of size $L_x \times L_y \times L_z = 10 \times 10 \times 10\,\mathrm{mm}^3$ was used, and each direction was uniformly divided by 512 grid points [28]. The grid size is $\delta_x \approx 20\,\mu$m, which ensures around 30 grid points within the smallest chemical thickness of the methane–air combustion. A modified chemical mechanism, MS-58, based on the Smooke and Giovangigli scheme [30] and improved to include OH^* chemistry [31], was utilized in the DNS. For a thorough description and verification of this mechanism, the reader is referred to Ref. [28]. The DNS was conducted using the SENGA2 source code [32] with the timestep of $\delta t = 1\,$ns. Around 60 snapshots were collected over a period of half of the flow through time $\tau_f = L_x/U_{in}$, where $U_{in} = 20\,$m/s is the inflow bulk mean velocity, after allowing the initial transients to escape the computational domain. Further details such as numerical schemes, boundary and initial conditions are discussed in Ref. [28] and the interested readers are referred to that study.

Table 1: Initial conditions of the DNS case used as the ANN training dataset.

Case	Λ_0/l_Z	l_c/l_Z	$X_{O_2}^{max}$	Z_{st}	$\langle Z\rangle$	$\sigma_Z/\langle Z\rangle$	$\langle c\rangle$	$\sigma_c/\langle c\rangle$
AZ1	0.60	0.77	0.035	0.01	0.008	1.05	0.56	0.46

2.2. Input and output selection

The selection of input features and output targets for the ANN is based on the modelling of the presumed FDF. A variety of shapes for the marginal FDF can be specified using the mean and variance. This has been explored and discussed in past studies [4, 33]. The use of the β-function is more typical for the mixture fraction FDF, although this function has also been used for the reaction progress variable. Following the DNS study by Doan *et al.* [28], the progress variable is defined using the temperature as

$$c_T = \frac{T - T_u}{T_b(Z) - T_u}, \tag{1}$$

where $T_u = 1500$K is the initial temperature for the unburnt mixture, and the burnt mixture temperature, T_b, depends on the local mixture fraction. The value of T_b is computed as described in [29]. The mixture fraction is based on Bilger's definition [34]:

$$Z = \frac{\beta - \beta_{ox}}{\beta_f - \beta_{ox}}, \tag{2}$$

where the subscripts 'f' and 'ox' represent fuel and oxidizer streams, respectively, and $\beta = 2Y_C/W_C + 0.5Y_H/W_H - Y_O/W_O$ is determined by elemental mass fractions and the respective molecular weight of carbon, hydrogen, and oxygen in the mixture. Within the scope of LES, the ANN inputs are defined by the first and second moments of the filtered progress variable and mixture fraction: $\widetilde{c_T}$, $\widetilde{\sigma^2_{c_T}}$, $\widetilde{Z}$, and $\widetilde{\sigma^2_Z}$, where $\widetilde{\cdot}$ represents Favre filtered quantities.

The filtered progress variable, $\widetilde{c_T}$, is directly utilized as the first input for the ANN since the values in the domain of most flames vary between the whole range of c_T values, $0 \leq \widetilde{c_T} \leq 1$. On the other hand, the range of $\widetilde{Z}$

can be different from case to case. These values depend on the fuel and the local equivalence ratio in flames. The limits of $\widetilde{Z}$ for the MILD condition considered here roughly range between 10^{-3} and 2×10^{-2} as shown in Fig. 1a with a bias towards the lean side. A more symmetric distribution can be obtained by using:

$$\psi = \ln \frac{\widetilde{Z}}{Z_{st}}, \tag{3}$$

where Z_{st} is the stoichiometric mixture fraction, as shown in Fig. 1b. This kind of distribution is preferred for ANN. The other two input features are the segregation factors, $\widetilde{g}_c$ and $\widetilde{g}_Z$, where $\widetilde{g}_Z$ is defined as:

$$\widetilde{g}_Z = \frac{\widetilde{\sigma_Z^2}}{\widetilde{Z}(1-\widetilde{Z})}, \tag{4}$$

and $\widetilde{g}_c$ is defined in a similar way. The sub-grid Favre mixture fraction variance is $\widetilde{\sigma_Z^2}$. The ANN trained using $\widetilde{g}_c$ and $\widetilde{g}_Z$ performs better than the one trained using $\widetilde{\sigma_{c_T}^2}$ and $\widetilde{\sigma_Z^2}$.

It was shown by Chen *et al.* [27] that the progress variable and mixture fraction are correlated for the MILD combustion conditions of AZ1. However, the interest for this study is to test the ANN that is trained using the MILD combustion case AZ1 for different sets of turbulent premixed flames. Hence, the sub-grid fluctuations of c_T and Z are assumed to be statistically independent, and the joint PDF is $\widetilde{P}(c_T, Z) = \widetilde{P}(c_T) \times \widetilde{P}(Z)$ for the training stage. Thus, the two marginal FDFs, $\widetilde{P}(c_T)$ and $\widetilde{P}(Z)$, are the outputs of ANN. This allows for reductions in the computational costs for the training and testing stages. This simplification does not compromise the accuracy of ANN, as discussed in Section 3.1. Further improvements to the ANN for inferring joint FDF, as done by Chen *et al.* [27], will be explored in the

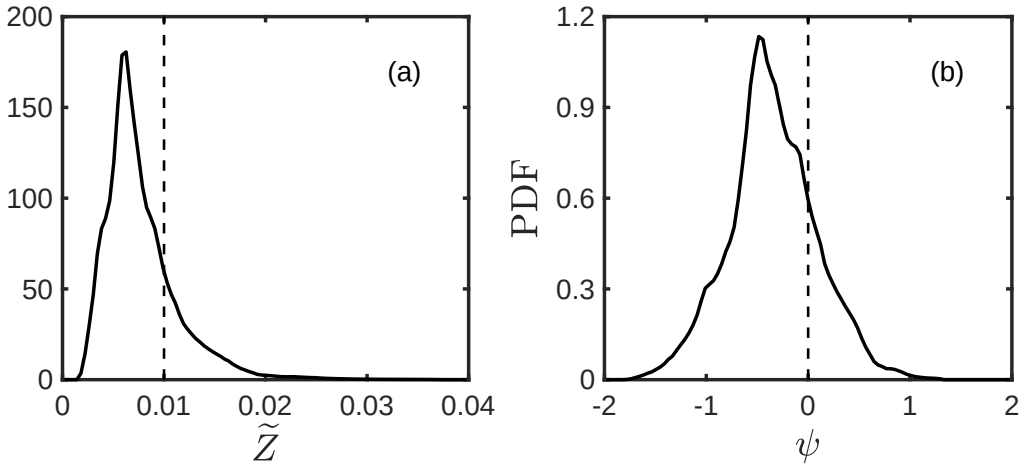

Fig. 1: PDF of (a) filtered mixture fraction $\widetilde{Z}$ and (b) scaled mixture fraction ψ for the training case AZ1. Solid line: PDF distribution; dashed line: stoichiometric value.

future.

2.3. Data extraction

A cubic box filter with size $\Delta = 80\delta x$, where δx is the DNS grid size, is used to filter the DNS fields. The width Δ equals the laminar thermal thickness under MILD condition, δ_{th}^{st}, i.e. $\Delta^+ = \Delta/\delta_{th}^{st} = 1$, following an earlier study [27]. The Favre filtered progress variable and its variance are calculated using

$$\widetilde{c_T}(\mathbf{x},t) = \frac{1}{\overline{\rho}(\mathbf{x},t)\Delta^3}\int_{\mathbf{x}-\frac{\Delta}{2}}^{\mathbf{x}+\frac{\Delta}{2}} \rho(\mathbf{x}',t)c_T(\mathbf{x}',t)\,d\mathbf{x}' \tag{5}$$

$$\widetilde{\sigma_{c_T}^2}(\mathbf{x},t) = \frac{1}{\overline{\rho}(\mathbf{x},t)\Delta^3}\int_{\mathbf{x}-\frac{\Delta}{2}}^{\mathbf{x}+\frac{\Delta}{2}} \rho(\mathbf{x}',t)\left[c_T(\mathbf{x}',t)-\widetilde{c_T}(\mathbf{x},t)\right]^2\,d\mathbf{x}' \tag{6}$$

where $\bar{\cdot}$ denotes filtered quantity without density weighting. The DNS and local filter sub-space are denoted using coordinate vectors, $\mathbf{x}$ and $\mathbf{x}'$, respec-

tively. The Favre filtered mixture fraction and its variance, $\widetilde{Z}$ and $\widetilde{\sigma_Z^2}$, are calculated using similar procedures. With these four quantities, the ANN input features, i.e., $\widetilde{c_T}$, $\widetilde{g}_c$, ψ, and $\widetilde{g}_Z$, are obtained using Eqs. (3) and (4).

The training targets are the two marginal FDFs. One is $\widetilde{P}(c_T)$ and is defined as

$$\widetilde{P}(\zeta; x, t) = \frac{1}{\overline{\rho}(x,t)} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \rho(x', t)\delta(\zeta - c_T(x', t))\, dx', \tag{7}$$

where ζ is the sample-space variable for c_T and $\delta(\cdot)$ is the Dirac delta function. The other marginal FDF for the mixture fraction, $\widetilde{P}(\psi)$, is obtained similarly. These FDFs are computed using 45 bins distributed uniformly within $0 \leq c_T \leq 1$ and 55 bins in ψ space, $-1.5 \leq \psi \leq 4.61$ based on extensive tests using different number of bins. Out of these 55 bins, 33 uniformly distributed bins are within $[-1, 0.3]$ since the majority of the MILD combustion data used for training is in this range, as indicated by the result shown in Fig. 1b. The interest here is to test the ANN trained on MILD combustion to various turbulent premixed flames. Hence, the range of scaled mixture fraction values is extended to 1.8, and the sub-range $[0.4, 1.8]$ is covered using 19 uniformly distributed bins. Reaction rate closures in the presumed PDF approach requires the FDF $\widetilde{P}(Z)$ which is related to $\widetilde{P}(\psi)$ through

$$\widetilde{P}(Z) = \frac{\widetilde{P}(\psi)}{Z}. \tag{8}$$

After obtaining one set of input-output pairs, the box filter is moved by a distance of $d = 20\delta x$ along one direction of the DNS domain. It should be noted that this distance should be larger than or equal to Δ to avoid overlaps. However, using such a choice will result in insufficient training

samples because of the limited number of DNS snapshots and the small physical size of the DNS domain. Hence, $d = 20\delta x$ is used for the current work along with three snapshots of DNS data. The total number of samples extracted is about 32000, which are used for training the ANN.

Overall, the input matrix with 31944 rows, representing the number of samples, and four columns, denoting the input features, is constructed. The output matrix has the same number of rows but 100 columns, where the first 45 contain $\widetilde{P}(c_T)$ values and the remaining 55 are $\widetilde{P}(\psi)$ values. For the input matrix, centring and scaling are used as effective methods to improve the performance of the machine learning (ML) algorithm. Furthermore, the potential influence of outliers of the sample on the training is controlled by removing the outliers by using principle component analysis (PCA) [35].

2.4. ANN setup

The input-output pair is connected by a feed-forward ANN, as shown in Fig. 2. The information is passed from the input layer to the subsequent layers and processed within neurons, represented by circles in Fig. 2. For the j^{th} neuron at the k^{th} layer, such a procedure can be expressed as

$$y_j^k = g^k \left(\sum_{i=1}^{N_i} W_{ij}^k x_i^{k-1} + b_j^k \right) = g^k a_j^k, \tag{9}$$

where y and x denote the output at k^{th} and input from the $(k-1)^{th}$ layer respectively and N_i denotes the total number of neurons at the $(k-1)^{th}$ layer. At each neuron, x is processed by a linear operation, a as shown in Fig. 2, using the weight connecting two continuous layers, W, and the bias term, b. Then the result is non-linearly processed by the activation function $g(\cdot)$. Before the final output layer, all other layers are fully connected since

the progress variable and the mixture fraction are correlated. Neurons at the output layer are divided into two groups, whereby 45 neurons are fully connected to half of the neurons in the penultimate layer, and the remaining 55 are connected to the other half. This ensures that $\widetilde{P}(c_T)$ and $\widetilde{P}(\psi)$ are retrieved from them respectively. This specific connection is made based on the assumption of statistical independence between the subgrid fluctuations of c_T and Z. This specific aspect halves the network computational cost compared to the full connection. At the output layer, the softmax function is utilized as the activation function, and the output from j^{th} neuron, y_j is:

$$y_j = \frac{\exp(a_j)}{\sum_{i=1}^{N} \exp(a_i)}, \tag{10}$$

where $N = 45$ for the outputs related to $\widetilde{P}(c_T)$ and 55 for $\widetilde{P}(\psi)$. Through this activation function, the result y_j is limited within [0, 1], which is the range for the FDF value. Each sum of $\widetilde{P}(c_T)$ or $\widetilde{P}(\psi)$ is unity, satisfying that the integral of FDF over sample space equals one. The difference between the output from ANN, y, and the expected value from DNS, $\hat{y}$, is measured using a binary cross-entropy loss function and back propagates towards the input layer. The weight at each layer is tuned by the gradient-descent method.

Among numerous factors affecting the ANN performance, the number of neurons in each layer, the number of layers, and the type of activation function at the hidden layers (between the input and output layer) are critical. Hence, the above three aspects are systematically varied for the 12 ANNs listed in Table 2 so that an optimised ANN can be found[2]. Every ANN has 32

[2]Techniques like pattern search algorithm may be employed to find an optimum ANN as has been done in past studies [19, 22], which will be explored in future investigations.

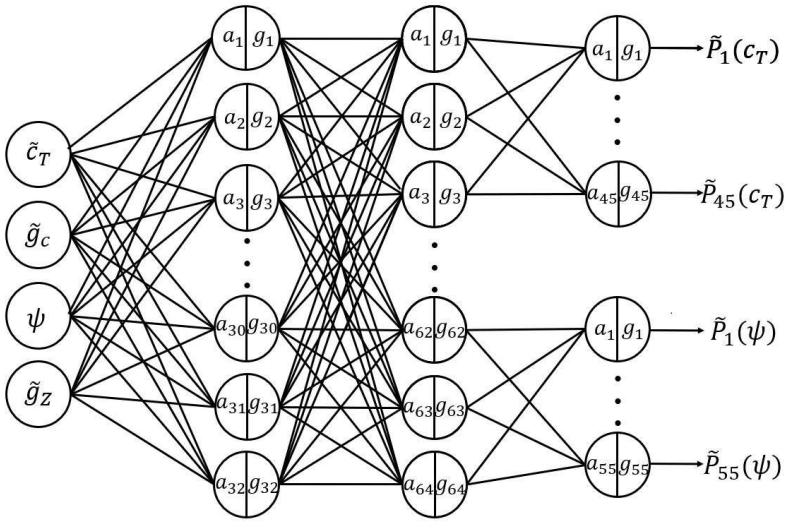

Fig. 2: Structure of the feed-forward ANN.

neurons in the first hidden layer and 64 neurons in the subsequent layers. The number of layers increases from 2 (cases 1-4) to 4 (cases 9-12), while for ANNs with the same number of layers, their activation functions are different. The leaky rectified linear unit (LeakyReLU) function, which is a common option for regression cases, is adopted for cases 1, 5, and 9. The hyperbolic tangent function and the sigmoid function are the activation functions for cases 2, 6 and 10, and 3, 7 and 11, respectively as listed in Table 2. The output is limited within $[-1, 1]$ in the former and $[0, 1]$ in the latter, which are similar to the range of the target data. Another activation function with a limited range, i.e., $[0, 1]$, is the hard sigmoid $S(x) = \max(0, \min(1, (x + 1)/2)$. Its performance is explored through the other 3 cases (4, 8, and 12).

Other hyperparameters within the ANN are optimised to the same value for all 12 cases. The whole training dataset is separated into several mini-batches of size 512 and batch normalization [36] is used. To reduce the

Table 2: Hidden layers attributes for the ANNs studied.

ANN	Number of neurons	Activation function
1	[32 64]	LeakyReLU
2	[32 64]	tanh
3	[32 64]	sigmoid
4	[32 64]	hard sigmoid
5	[32 64 64]	LeakyReLU
6	[32 64 64]	tanh
7	[32 64 64]	sigmoid
8	[32 64 64]	hard sigmoid
9	[32 64 64 64]	LeakyReLU
10	[32 64 64 64]	tanh
11	[32 64 64 64]	sigmoid
12	[32 64 64 64]	hard sigmoid

computational cost, an early stopping method is deployed so that the training would stop after 25 consecutive epochs that have not witnessed a loss function decrease. The Adam optimization method [37] is used to adjust the learning rate to minimise the loss function smoothly. In order to prevent overfitting, the L2 norm is added as a regularization method [38]. All ANNs are trained over a maximum number of 1000 epochs on a local workstation with Intel Xeon E5-2630 v3 CPU. As an *a priori* study, both training and testing

activities are conducted using the TensorFlow library [39]. Regarding the computational cost, the data extraction process is the most time-consuming. The preprocessing of each snapshot of MILD combustion requires around 30 minutes, while the training takes about 5-10 minutes for each ANN.

2.5. Test cases

In the present study, three H_2–air and two CH_4–air combustion cases are considered. The stoichiometric H_2–air cases are (i) a freely-propagating statistically planar flame, R97CR [40–44], (ii) a turbulent V-flame anchored by a hot rod, V97 [45], and (iii) an adiabatic swirl flame with swirl number of 1.2, S12 [46, 47], which are listed in Table 3. Two CH_4–air lean combustion cases are a twin V-flame, TV06 [48], and a stagnation flame, SC [49]. Table 3 lists the thermo-chemical and -physical conditions of these cases. Here, u' is the inlet turbulence intensity with an integral length scale of l_0, while s_L and δ_{th} are the unstrained planar laminar flame speed and its thermal thickness, respectively. The mixture fraction values are also listed. The Damköhler number is defined as $Da = (l_0/\delta_{th})/(u'/s_L)$, and the Karlovitz number is defined as $Ka = (u'/s_L)^{3/2}(l_0/\delta_{th})^{-1/2}$. For the swirl flame (S12), the turbulence integral length scale is not listed in the corresponding studies [46, 47]. Thus, the value of Da and Ka based on inlet conditions are not available, while their ranges are estimated approximately from the regime diagram shown in Ref. [46].

Table 3 also lists the DNS domain sizes in mm and the number of grid points used in each direction. To be consistent with the training procedure, a box filter with size $\Delta = \delta_{th}$ (i.e., $\Delta^+ = 1$) is used to extract test data. Table 4 lists the pertinent details of the testing data extracted using a single

Table 3: Numerical conditions of test cases for ANN.

Case	Z	u'/s_L	l_0/δ_{th}	Da	Ka	N_x, N_y, N_z	L_x, L_y, L_z [mm]	Ref.
R97CR	0.0282	2.88	3.22	1.12	2.72	1025, 256, 256	21.6, 10.8, 10.8	[40–44]
V97	0.0282	6.0	1.5	0.24	12	769, 385, 385	10.0, 5.0, 5.0	[45]
S12	0.0282	1.245	-	< 1	≥ 1	769, 513, 513	15.0, 10.0, 10.0	[46, 47]
TV06	0.0346	6.0	3.1	0.51	8.35	672, 336, 588	12.8, 6.4, 11.2	[48]
SC	0.0346	12.5	0.9	0.08	46.6	384, 256, 256	7.5, 5.0, 5.0	[49]

snapshot from each case listed in Table 3. For the stoichiometric hydrogen–air flames, $Z = Z_{st}$, which implies that the input feature $\psi = 0$, whereas $\psi = -0.4602$ for the methane-air flames. Since these cases are premixed, there is no mixture fraction variation in space and time. Hence, the input feature $\widetilde{g}_Z$ is zero.

The progress variable defined in Eq. (1) is used for the H_2–air flames with $T_u = 700$ K and $T_b = 2412$ K, which are obtained from the corresponding 1D laminar flames. Since the DNS of CH_4–air flames used a single-step irreversible global reaction chemistry, the progress variable is defined as $c = Y_p$, where Y_p is the product mass fraction, following [48, 49]. Since the combustion is adiabatic and the CH_4–air mixture has a Lewis number close

Table 4: Summary of the testing dataset.

Case	Δ^+	Δ/δ_x	sample size	Mixture
R97CR	1	24	4200	H_2-air
V97	1	14	39366	H_2-air
S12	1	26	6525	H_2-air
TV06	1	22	11700	CH_4-air
SC	1	22	2057	CH_4-air

to unity, the progress variable is $c = c_T$.

3. Results

3.1. Validation and optimization

The 31944 samples from three snapshots of the DNS case AZ1 are split into two sets having 80% and 20% of the samples. These two sets are used for training and validation, respectively. The former is used to train the ANNs listed in Table 2, and the latter is used to control the training progress, including adjusting the learning rate and determining the convergence. Once all ANNs are trained, they are tested on the same training dataset. The Jensen-Shannon divergence (JSD) [50] between their outputs and DNS values is calculated and compared. The JSD is expressed as

$$\mathrm{JSD}(Q_1||Q_2) = \frac{1}{2}\sum_{n=1}^{N}\left(Q_1(n)\ln\frac{Q_1(n)}{Q_2(n)} + Q_2(n)\ln\frac{Q_2(n)}{Q_1(n)}\right), \qquad (11)$$

where Q_1 and Q_2 are probability distributions with size N. The JSD is mathematically bounded between 0 and $\ln(2)$, with 0 indicating $Q_1 = Q_2$.

Hence, a lower JSD value implies a more similar distribution between Q_1 and Q_2, namely, a better FDF prediction by the ANN in this work. A JSD value is obtained for each box filter. The average of JSD values for the two marginal FDFs, $\widetilde{P}(c_T)$ and $\widetilde{P}(Z)$, between ANN and DNS are displayed in Fig. 3 for all the 12 ANNs considered. It is noted that the marginal FDF is conditioned by (c_T, Z) to eliminate the randomness in the FDF [27], and it is denoted using the symbol $\langle\cdot\rangle$. Overall, all of the ANNs show low values of JSD, which are $\overline{\langle JSD_c\rangle} \leq 0.035$ and $\overline{\langle JSD_Z\rangle} \leq 0.06$. These mean JSDs are very similar to JSDs reported in [27] for the cases where c and Z are taken to be statistically independent. The corresponding mean JSDs for $\widetilde{P}(c)$ and $\widetilde{P}(Z)$ reported in [27] are about 0.03 and 0.045, respectively.

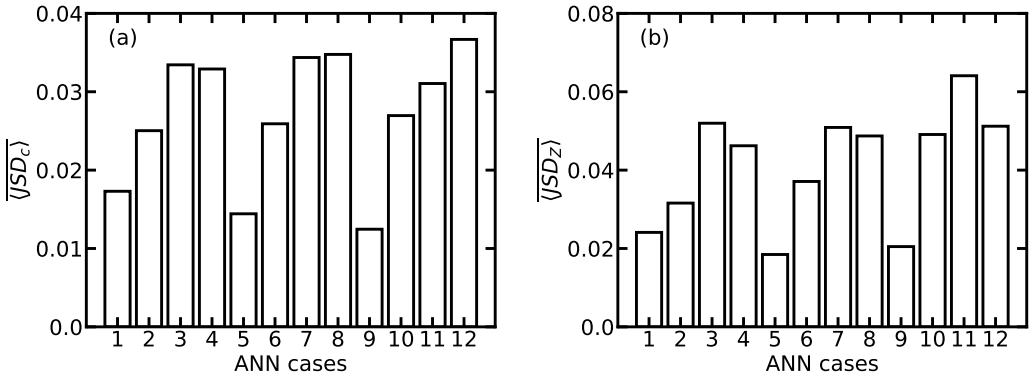

Fig. 3: Averaged JSD values for the marginal FDFs of (a) progress variable, and (b) mixture fraction provided by the ANN cases listed in Table 2.

Figure 3 illustrates the superior performance of LeakyReLU (case 1, 5 and 9) compared to other activation functions. For the shallowest layout, the ANN with LeakyReLU (case 1) exhibits the lowest JSD value for both the progress variable and mixture fraction. When deepening the ANN,

LeakyReLU is the only activation function showing benefits. One can say that the LeakyReLU is the best activation function for the current modelling task. Among LeakyReLU cases (1, 5, and 9), a continuous drop in $\overline{\langle JSD_c \rangle}$ is observed in Fig. 3a. A similar drop is also noted for $\overline{\langle JSD_Z \rangle}$ but using four hidden layers (case 9) increases the JSD slightly as seen in Fig. 3b. Based on this result, the optimised ANN, case 9 in Table 2, is tested for various turbulent premixed flames, which differ from the MILD combustion training case.

3.2. FDF prediction

Figure 4 shows the marginal FDFs, $\langle \widetilde{P}(c) \rangle$ for all testing cases. The brackets $\langle \cdot \rangle$ denote the quantity conditioned on $\widetilde{c}$ and $\widetilde{g}_c$ and ensemble-averaged over the whole computational domain. The corresponding values of $\widetilde{c}$ and $\widetilde{g}_c$ are displayed in each plot. To demonstrate the potential of replacing the conventional presumed FDF model with the ANN, the FDF predicted by ANN is compared with the same quantity predicted using the β-PDF method. This presumed distribution is obtained through

$$\widetilde{P}(c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c^{a-1}(1-c)^{b-1}, \tag{12}$$

where $a = \widetilde{c}(1/\widetilde{g}_c - 1)$, $b = (1-\widetilde{c})(1/\widetilde{g}_c - 1)$ and Γ is the *gamma* function. The corresponding β function FDFs are shown in Figure 4 as well. Overall, the results predicted by ANN agree fairly well with the DNS results and are comparable with the β function FDFs. For the premixed hydrogen–air flames (Fig. 4a), the marginal FDFs shown in the top row of Fig. 4 are monomodal and are well predicted by the β-PDF method. The ANN model gives a similar value but the peak FDF is underpredicted. This behaviour is

emerging because the training data from the MILD combustion case have a broader distribution in the progress variable space. Meanwhile, a slight right skew is observed for cases V97 and S12 (see last two subfigures in Fig. 4a). For premixed methane–air combustion, the marginal FDFs show different shapes. Figure 4b illustrates a wide quasi-plateau for TV06, and exhibits a single peak accompanied by a plateau towards the unburnt side for SC. These shapes are predicted by the ANN quite well, especially for the SC case, where both plateau and peak are captured and a lower skewness discrepancy can be appreciated. By contrast, the presumed β-PDF method does not capture these special shapes.

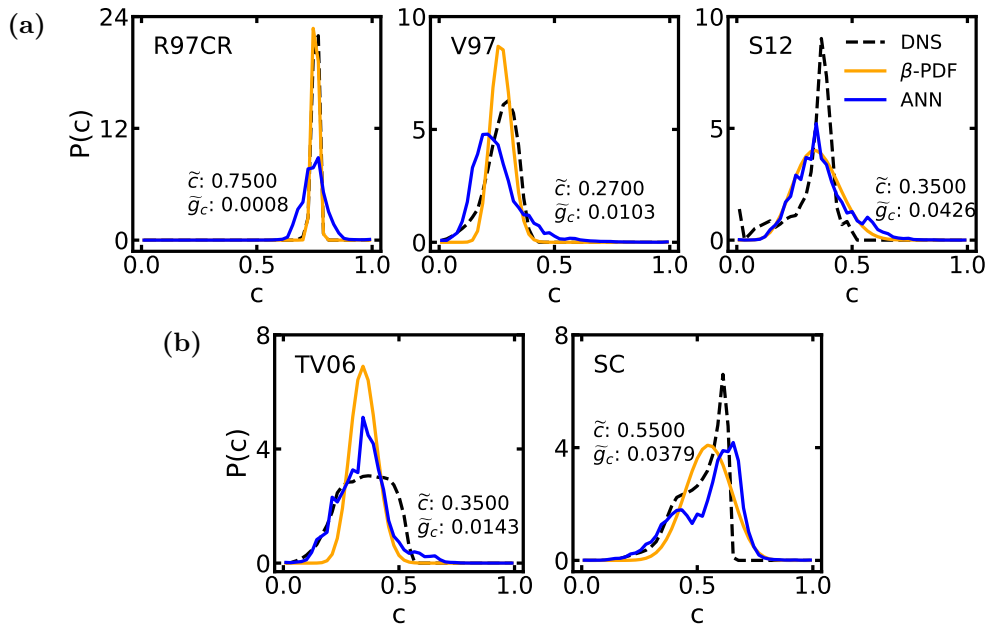

Fig. 4: Comparison of the conditional marginal FDFs of progress variable inferred using the ANN (solid blue lines) with the β-PDF method (solid orange lines) and the DNS results (dashed lines) for (a) H_2–air flames, and (b) CH_4–air flames.

To further assess the predictive ability of the trained ANN on different testing cases, the JSD between the FDFs from DNS and the ANN is computed for the five premixed flames. The PDFs of JSD are illustrated in Fig. 5 with the mean values of JSDs, as well as the results from the β-PDF method. For the filter size $\Delta^+ = 1$ (solid lines), the peak and widest area of PDF plots predicted by ANN for each case are mainly located below 0.3, implying a good performance in extrapolation. As a result of underprediction and right skewness seen in Fig. 4a, the ANN's JSD values for the three hydrogen–air flames are generally not very low. The mean values are around 0.25 (see Fig. 5a), compared with around 0.08 for the β-PDF method. Regarding the methane–air flames, the ANN accuracy improves and becomes higher than the β-PDF method. The JSD values cluster closer to zero with lower mean values of JSDs, as shown in Fig. 5b. These satisfactory predictions prove the robustness of the ANN for different flames. In terms of the robustness for different filter sizes, $\Delta^+ = 2$ is also used to filter the DNS data. The JSD values for the new ANN tested on different premixed flames are shown in Fig. 5 by using the dashed line. The different sensitivity to the filter size for different fuels will be further investigated in future.

3.3. Source term prediction

Through FDF modelling, the filtered reaction rate $\overline{\dot{\omega}}_c$ is obtained as

$$\overline{\dot{\omega}}_c(x,t) = \overline{\rho}(x,t) \int_0^1 \frac{\dot{\omega}_c(\zeta)}{\rho(\zeta)} \widetilde{P}(\zeta;x,t)\, d\zeta \tag{13}$$

where ζ is the sample space variable for the progress variable, as defined in Eq. (7), $\dot{\omega}_c(\zeta)$ and $\rho(\zeta)$ are the flamelet reaction rate and density, respectively, obtained from 1D unstrained premixed laminar flame. The flamelets

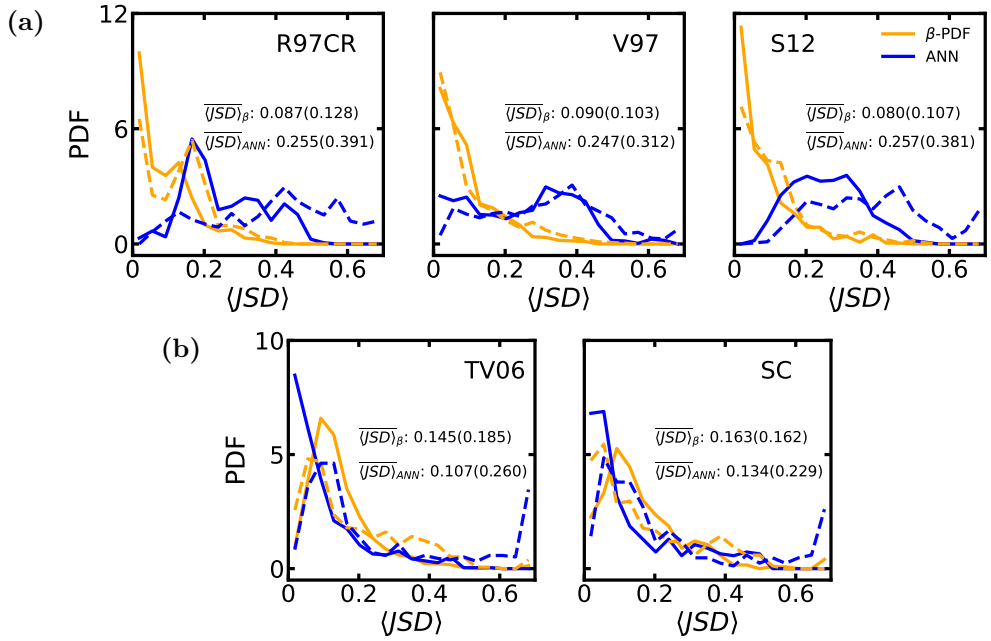

Fig. 5: PDF of JSD for conditional FDF from the ANN (blue) and β-PDF method (orange) for cases (a) H_2–air and (b) CH_4–air. The solid and dashed lines correspond the filter size $\Delta^+ = 1$ and $\Delta^+ = 2$ respectively. The mean JSD values for $\Delta^+ = 1$ ($\Delta^+ = 2$) are also listed.

are computed using Cantera [51] with the GRI-Mech 3.0 [52] chemical mechanism. Since c is defined differently for H_2–air and CH_4–air flames, the laminar reaction rate $\dot{\omega}_c(\zeta)$ is calculated accordingly. For cases R97CR, V97, and S12, $\dot{\omega}_c = \dot{q}/(c_p(T_b - T_u))$, where $\dot{q}$ and c_p represent the heat release rate per volume and the specific heat capacity at constant pressure for the mixture respectively. For cases TV06 and SC, $\dot{\omega}_c$ is based on the reaction rates for CO_2 and CO, i.e. $\dot{\omega}_c = (\dot{\omega}_{CO} + \dot{\omega}_{CO_2})/(Y_{CO} + Y_{CO_2})_b$, where the subscript '$b$' denotes the burnt mixture, and Y_{CO} and Y_{CO_2} are the mass fractions of CO and CO_2, respectively. Once the laminar thermo-chemical quantities are

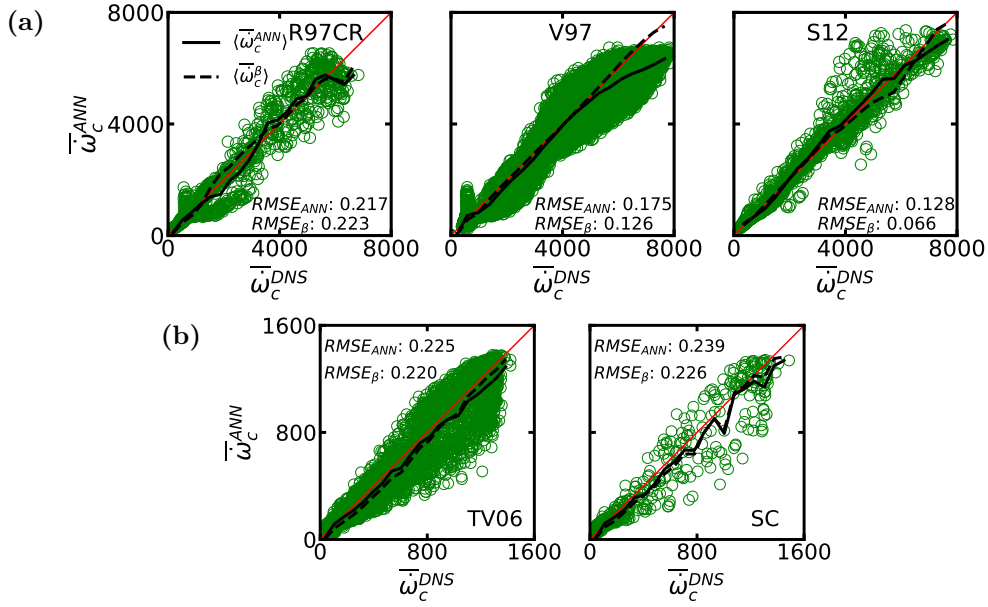

Fig. 6: Scatter plot of progress variable source terms from ANN 9 in Table 2 and DNS tested on (a) H_2–air, and (b) CH_4–air combustion cases. Solid and dashed lines correspond to the conditional averages from the ANN and β-PDF respectively.

obtained, the marginal FDFs, $\widetilde{P}(\zeta)$, from the ANN are used in Eq. (13) to calculate the filtered reaction rate.

The source terms $\overline{\dot{\omega}}_c$ obtained using the ANN case 9 of Table 2 and extracted from the DNS are compared in the scatter plots of Fig. 6. It can be observed that the ANN provides good results since most predictions are clustered around the diagonal lines. Although the FDFs predicted by the ANN had some minor differences for the H_2–air flames, the reaction rates obtained using this ANN are reasonable, as shown in the top row of Fig. 6, while a slight underprediction can be noticed for the higher values of the source term. The predictions for the CH_4–air flames are symmetrically

located around the diagonal lines, indicating unbiased predictions. In terms of the conditional averaging quantity, the ANN shows a very good prediction for all cases. The conditional values are situated close to the diagonal lines (shown by the black solid lines in Fig.6), except for a slight deviation at the regions of the H_2–air flames having high source term values. The source terms obtained through β-PDF are also conditioned and depicted by black dashed lines in Fig. 6, and the ANN results show an excellent agreement with the β-PDF model, especially for CH_4–air combustion where no clear difference between the two methods can be observed.

A quantitative comparison between different models is computed using the root-mean-square error (RMSE) as

$$RMSE = \sqrt{\frac{1}{N}\left(\frac{\overline{\dot{\omega}}_c^{ANN} - \overline{\dot{\omega}}_c^{DNS}}{\overline{\dot{\omega}}_c^{DNS}}\right)^2}, \tag{14}$$

where N is the data size. The RMSE for the source term modelled by the β-PDF method is calculated similarly. It is noted that only samples having $\overline{\dot{\omega}}_c^{DNS} \geq 0.1\max(\overline{\dot{\omega}}_c^{DNS})$ are considered for this error estimates. RMSE values for these two models are included in each frame of Fig. 6. It is observed that they are comparable to each other for all the testing cases implying that the predictive ability is similar for these two models although the ANN has higher RMSE values particularly for hydrogen–air flames.

4. Further improvement

Training a pre-trained model using data from different but relevant cases is used in the machine learning community typically to extend the scope of application. This concept, namely fine-tuned training [53], is used in the

present study to expand the usability of the current ANN, which is influenced by the training data set in terms of flame type, combustion behaviour, flammability ranges, etc., as discussed in the previous section. One premixed flame from the testing cases is added to the training set. The same strategy described in Section 2.3 is used to extract data, while more snapshots are considered to keep the amount of data from the premixed flame comparable to that from the MILD combustion. Five ANNs, N1 to N5, having the same structure as the ANN case 9 of Table 2 are trained on the expanded training datasets, composed of the MILD and only one premixed case as listed in Table 5. The data size can also be increased by adding two cases at a time but this limits the testing data within the limited set available for this study. Hence this is not attempted.

Table 5: Summary of additional data added to the original MILD combustion training dataset.

New ANN	Additional training case	Snapshots	Sample size
N1	R97CR	8	37800
N2	V97	1	39366
N3	S12	5	32625
N4	TV06	3	34944
N5	SC	17	34969

Since one premixed case is used for the training, the new ANNs N1 to N5 are tested on the remaining four premixed cases. The averaged JSD values for the FDFs are compared in Fig. 7. The JSD value provided by

the previously trained ANN, discussed in Section 3.2, is represented by a dashed line. A significant improvement is observed when the premixed flame used for testing is similar to the one used for training. Regarding the H_2–air flames, the new ANNs trained with additional data from hydrogen–air flames (N1-N3 in Table 5) show a clear decrease from the original JSD value, as seen by the first three subfigures in Fig. 7. By contrast, the ANNs trained with data from premixed methane–air flames (N4 and N5) show an increase in the JSD values, especially when testing on the S12 flame. For CH_4–air flames, ANNs N4 and N5 provide an excellent prediction with very low JSD values, as seen in Figs. 7d and 7e, while the performances of the ANNs trained with hydrogen–air combustion data slightly worsen.

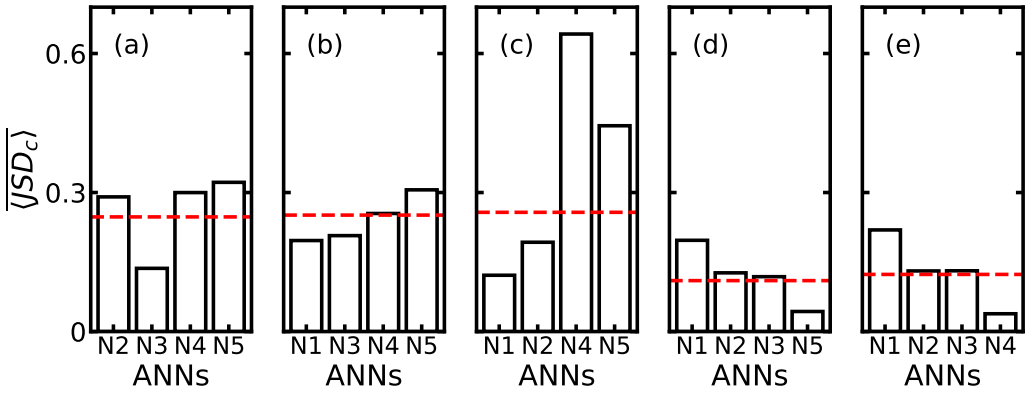

Fig. 7: Averaged JSD values for the marginal FDFs of progress variable from newly trained ANNs (on extended training datasets) for various cases: (a) R97CR, (b) V97, (c) S12, (d) TV06, and (e) SC. The averaged JSD value provided by ANN case 9 is depicted by a red dashed line.

Accordingly, the ANNs N1, N2, and N3, listed in Table 5, are used to calculate the filtered reaction rate for premixed H_2–air flames. The ANNs N4

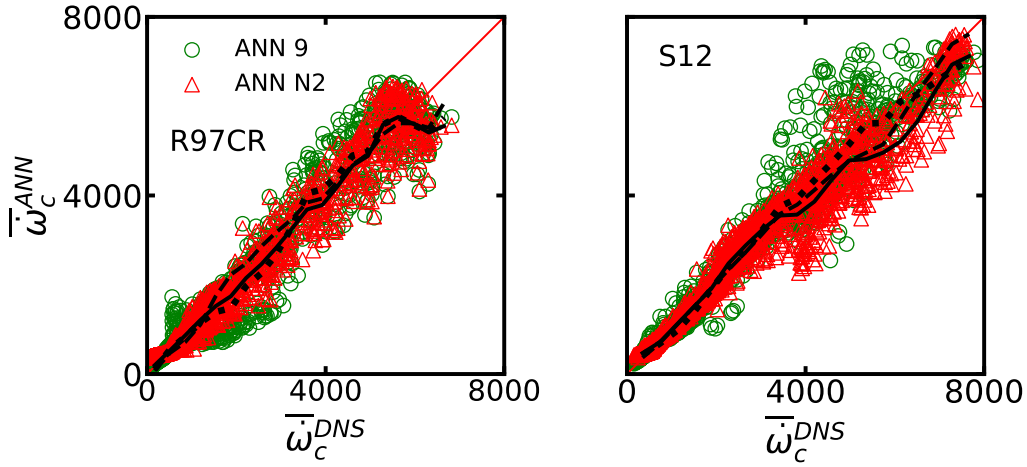

Fig. 8: Scatter plot of progress variable source terms from ANN N2 in Table 5 and DNS (red triangles), and ANN 9 in Table 2 and DNS (green circles), for different H_2–air combustion cases. Solid black line: $\langle\overline{\dot{\omega}}_c^{ANN}\rangle$; dotted black line: $\langle\overline{\dot{\omega}}_{c,0}^{ANN}\rangle$; dashed black line: $\langle\overline{\dot{\omega}}_c^{\beta}\rangle$. Subscript '0' denotes the ANN trained with data from the MILD combustion only (ANN 9 in Table 2).

and N5 are used for premixed CH_4–air flames. Among these two categories, the ANN modelling performance is similar. Hence, for brevity, only the results of ANNs N2 and N4, which are the most accurate in an overall sense, are discussed in this section.

Figure 8 shows the scatter plot of the filtered reaction rate calculated using the ANN N2 and the DNS values for the two cases, R97CR and S12. The results from the ANN 9 (see Table 2) trained using the MILD combustion data only are also shown. Compared with previous results, the inferences from the new ANN are more clustered along the diagonal, implying a closer agreement to the target values. This improvement is achieved by including the V97 data in the training set. The averaged data conditioned on DNS

quantities for the ANN N2, (black solid line), ANN 9 (black dotted line), and β-PDF method (black dashed line) are also shown in Fig. 8. By comparing different models, the ANN trained using the extended set matches the β-PDF better than the ANN trained with the MILD combustion alone. For the flame case R97CR, the conditional averaged quantity modelled by the updated ANN shows a quasi-diagonal trend. However, the underprediction at the high source terms region is still noticeable and occurs for a wider range after adding new data into the training dataset when testing the ANN, N2, on the flame S12. The N4 ANN trained with additional data from the CH_4–air flame TV06 exhibits a remarkable performance on the stagnation flame (SC) as shown in Fig. 9. The local and conditional values agree well with the DNS results, showing a considerable improvement with respect to the ANN trained with data from the MILD combustion only. Compared to the β-PDF model, smooth predictions without any fluctuations across the whole range of values can be observed. Therefore, it is reasonable to say that the updated ANN N4 has a superior predictivity than the conventional models for the premixed methane-air combustion cases considered here.

The RMSE values for the models shown in Figs. 8 and 9 are listed in Table 6. A noticeable drop in RMSE for all testing cases is seen after training the ANN with extended datasets. The RMSE values provided by the updated ANN are lower than the ones for the β-PDF approach for cases R97CR and SC. Conversely, the RMSE of the ANN N2 is slightly higher than that for the β-PDF for the case S12 in Table 6, which is due to the underestimation of the conditional filtered reaction rate, shown in Fig. 8. This may suggest an insufficient heterogeneity of the training dataset for premixed H_2–air flames.

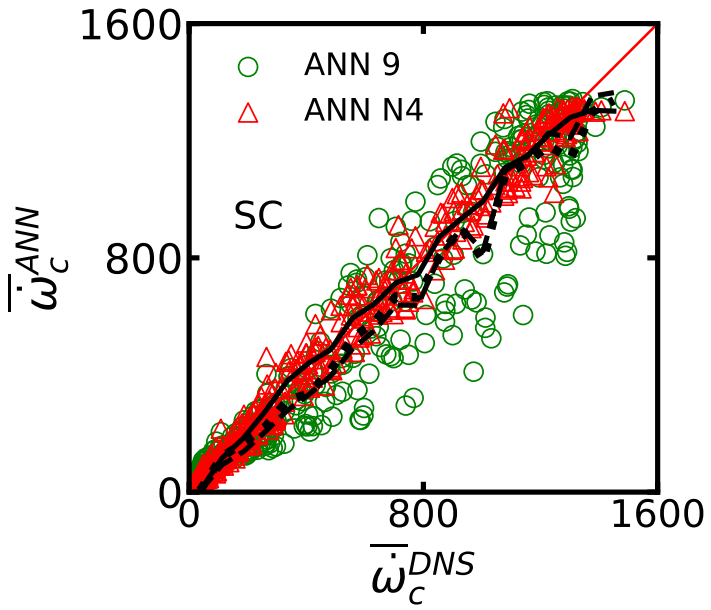

Fig. 9: Scatter plot of reaction rate source term from ANN N4 in Table 5 and DNS for premixed CH_4–air combustion case SC. Solid black line: $\langle\overline{\dot{\omega}}_c^{ANN}\rangle$; dotted black line: $\langle\overline{\dot{\omega}}_{c,0}^{ANN}\rangle$; dashed black line: $\langle\overline{\dot{\omega}}_c^{\beta}\rangle$. Subscript '0' denotes the ANN trained with data from the MILD combustion only (ANN 9 in Table. 2).

Extending the scope of conditions covered by the training data is an effective way to improve the ANN performance. Further refinement of the ANN to improve the inference of FDF and reaction rate for hydrogen–air flames will be explored in the future.

5. Conclusions

An artificial neural network (ANN) trained on DNS data of methane–air moderate or intense low-oxygen dilution (MILD) combustion is used to infer the filtered density functions (FDFs) of the mixture fraction and progress

Table 6: RMSE of modelled progress variable source terms for different modelling methods. The superscript 0 denotes the ANN 9 in Table 2.

Testing case	RMSE^0_{ANN}	RMSE_{ANN}	RMSE_β
R97CR	0.2166	0.1553	0.2228
S12	0.1280	0.1007	0.0657
SC	0.2547	0.1356	0.2258

variable for entirely different combustion cases and conditions. The optimum network structure and activation function are investigated thoroughly at first. An ANN with four hidden layers and the leaky rectified linear unit (LeakyReLU) activation function yields the best accuracy for inferring FDFs in MILD combustion. This ANN is then tested on 5 different DNS datasets of premixed flames to verify the generalisation capability of the ANN. It is observed that this ANN is able to infer the FDF quite well for all the flames of methane–air mixtures studied here. However, some differences between inferred and DNS FDFs are observed for premixed hydrogen–air flames.

The inferred FDFs are then used to compute the filtered reaction rate of a reaction progress variable using the presumed FDF approach. The filtered reaction rate is compared to the DNS result and that obtained using β-FDF. A fairly good agreement between the DNS and the ANN results is observed for each of the five testing cases. However, some underprediction of the filtered source term for the hydrogen–air flames can be noticed in the region of high reaction rates. The conditionally averaged filtered reaction rate calculated using the ANN and β-FDF agree well, especially for methane–air

flames, while small differences are observed for hydrogen–air flames.

An improvement to the ANN performance is also investigated by combining one premixed DNS case with the MILD dataset. This gives five new ANNs as listed in Table 5, which are also tested by excluding the premixed case used for training. An increase in the accuracy of the marginal FDF predictions is observed for these revised ANNs, ensuring that a limited scope of the training data is sufficient for good results in a broad range of testing cases. Two ANNs are finally selected; one is trained on the data from CH_4–air MILD combustion and a H_2–air V-shaped flame, and the other is trained using the MILD combustion and CH_4–air twin V-flame. They are employed to model the filtered progress variable source term for hydrogen–air flames and methane–air flames, respectively. An excellent agreement with DNS results is observed, although a slight underestimation for the hydrogen–air flames is still seen. Compared to the results obtained using the classical β-FDF, ANN predictions are more accurate, as indicated by lower RMSE values. This opens up potential opportunities to employ ANN in LES of reacting flows, which will be explored in a future study. Furthermore, assessing the predictive ability of the trained ANNs for partially premixed combustion, involving mixture fraction variations, of hydrogen and methane and for turbulent combustion of other fuels will be interesting and useful extensions of this study. Also, the sensitivity to including non-premixed DNS data in the training set needs to be investigated.

6. Acknowledgements

S.I. acknowledges the financial support of the Fonds National de la Recherche Scientifique (FRS-FNRS). J.C.M. and N.S. acknowledge the financial support from Mitsubishi Heavy Industries, Ltd., Takasago, Japan. N.S. acknowledges the support of EPSRC through grant number EP/S025650/1. Z.X.C. acknowledges the support of the National Science Foundation of China (Grant No. 52276096 and No. 92270203). The authors thank Professor Tanahashi for pre-processed DNS data for the premixed flames.

References

[1] P. Sagaut, Large eddy simulation for incompressible flows, Scientific Computation, Springer Berlin, Heidelberg, Berlin/Heidelberg, 3 edition, 2006.

[2] S. Menon, C. Fureby, Encyclopedia of aerospace engineering, John Wiley Sons, Ltd, Chichester, UK, 2010.

[3] S. B. Pope, Turbulent flows, Cambridge University Press, Cambridge, UK, 2000.

[4] T. Poinsot, D. Veynante, Theoretical and numerical combustion, R.T. Edwards, Philadelphia, PA, 2 edition, 2005.

[5] F. Christo, A. Masri, E. Nebot, T. Turanyi, Utilising artificial neural network and repro-modelling in turbulent combustion, Proc. ICNN'95 - Int. Conf. Neural Networks, volume 2, IEEE, 1995, pp. 911–916.

[6] F. Christo, Artificial neural network implementation of chemistry with pdf simulation of H2/CO2 flames, Combust. Flame 106 (1996) 406–427.

[7] F. Christo, A. Masri, E. Nebot, S. Pope, An integrated PDF/neural network approach for simulating turbulent reacting systems, Symp. Combust. 26 (1996) 43–48.

[8] J. Blasco, N. Fueyo, C. Dopazo, J. Ballester, Modelling the temporal evolution of a reduced combustion chemical system with an artificial neural network, Combust. Flame 113 (1998) 38–52.

[9] J. Blasco, N. Fueyo, J. Larroya, C. Dopazo, Y.-J. Chen, A single-step time-integrator of a methane–air chemical system using artificial neural networks, Comput. Chem. Eng. 23 (1999) 1127–1133.

[10] B. A. Sen, S. Menon, Linear eddy mixing based tabulation and artificial neural networks for large eddy simulations of turbulent flames, Combust. Flame 157 (2010) 62–74.

[11] B. A. Sen, S. Menon, Turbulent premixed flame modeling using artificial neural networks based chemical kinetics, Proc. Combust. Inst. 32 (2009) 1605–1611.

[12] B. A. Sen, E. R. Hawkes, S. Menon, Large eddy simulation of extinction and reignition with artificial neural networks based chemical kinetics, Combust. Flame 157 (2010) 566–578.

[13] A. Chatzopoulos, S. Rigopoulos, A chemistry tabulation approach via Rate-Controlled Constrained Equilibrium (RCCE) and Artificial

Neural Networks (ANNs), with application to turbulent non-premixed CH4/H2/N2 flames, Proc. Combust. Inst. 34 (2013) 1465–1473.

[14] L. L. Franke, A. K. Chatzopoulos, S. Rigopoulos, Tabulation of combustion chemistry via Artificial Neural Networks (ANNs): Methodology and application to LES-PDF simulation of Sydney flame L, Combust. Flame 185 (2017) 245–260.

[15] K. Wan, C. Barnaud, L. Vervisch, P. Domingo, Chemistry reduction using machine learning trained from non-premixed micro-mixing modeling: application to DNS of a syngas turbulent oxy-flame with side-wall effects, Combust. Flame 220 (2020) 119–129.

[16] F. Flemming, A. Sadiki, J. Janicka, LES using artificial neural networks for chemistry representation, Prog. Comput. Fluid Dyn. An Int. J. 5 (2005) 375.

[17] A. Kempf, F. Flemming, J. Janicka, Investigation of lengthscales, scalar dissipation, and flame orientation in a piloted diffusion flame by LES, Proc. Combust. Inst. 30 (2005) 557–565.

[18] M. Emami, A. Eshghinejad Fard, Laminar flamelet modeling of a turbulent CH4/H2/N2 jet diffusion flame using artificial neural networks, Appl. Math. Model. 36 (2012) 2082–2093.

[19] M. Ihme, A. L. Marsden, H. Pitsch, Generation of Optimal Artificial Neural Networks Using a Pattern Search Algorithm: Application to Approximation of Chemical Systems, Neural Comput. 20 (2008) 573–601.

[20] M. Ihme, C. Schmitt, H. Pitsch, Optimal artificial neural networks and tabulation methods for chemistry representation in LES of a bluff-body swirl-stabilized flame, Proc. Combust. Inst. 32 (2009) 1527–1535.

[21] O. Owoyele, P. Kundu, M. M. Ameen, T. Echekki, S. Som, Application of deep artificial neural networks to multi-dimensional flamelet libraries and spray flames, Int. J. Engine Res. 21 (2020) 151–168.

[22] R. Ranade, G. Li, S. Li, T. Echekki, An Efficient Machine-Learning Approach for PDF Tabulation in Turbulent Combustion Closure, Combust. Sci. Technol. 193 (2021) 1258–1277.

[23] M. Ihme, W. T. Chung, A. A. Mishra, Combustion machine learning: Principles, progress and prospects, Prog. Energy Combust. Sci. 91 (2022) 101010.

[24] N. Swaminathan, A. Parente, Machine Learning and Its Application to Reacting Flows, Springer International Publishing, Cham, 2023.

[25] M. T. Henry de Frahan, S. Yellapantula, R. King, M. S. Day, R. W. Grout, Deep learning for presumed probability density function models, Combust. Flame 208 (2019) 436–450.

[26] S. Yao, B. Wang, A. Kronenburg, O. T. Stein, Modeling of sub-grid conditional mixing statistics in turbulent sprays using machine learning methods, Phys. Fluids 32 (2020) 115124.

[27] Z. X. Chen, S. Iavarone, G. Ghiasi, V. Kannan, G. D'Alessio, A. Parente, N. Swaminathan, Application of machine learning for filtered density

function closure in MILD combustion, Combust. Flame 225 (2021) 160–179.

[28] N. A. K. Doan, N. Swaminathan, Y. Minamoto, DNS of MILD combustion with mixture fraction variations, Combust. Flame 189 (2018) 173–189.

[29] Y. Minamoto, N. Swaminathan, Scalar gradient behaviour in MILD combustion, Combust. Flame 161 (2014) 1063–1075.

[30] M. D. Smoke, V. Giovangigli, Formulation of the premixed and nonpremixed test problems, Reduc. Kinet. Mech. Asymptot. Approx. Methane-Air Flames, Springer Berlin, Heidelberg, 1991, pp. 1–28.

[31] T. Kathrotia, U. Riedel, A. Seipel, K. Moshammer, A. Brockhinke, Experimental and numerical study of chemiluminescent species in low-pressure flames, Appl. Phys. B 107 (2012) 571–584.

[32] R. S. Cant, SENGA2 user guide, Technical Report, Cambridge University Engineering Department, 2012.

[33] H. Pitsch, Large-eddy simulation of turbulent combution, Annu. Rev. Fluid Mech. 38 (2006) 453–482.

[34] R. Bilger, S. Stårner, R. Kee, On reduced mechanisms for methane-air combustion in nonpremixed flames, Combust. Flame 80 (1990) 135–149.

[35] I. T. Jolliffe, Principal component analysis, Springer Series in Statistics, Springer New York, NY, New York, 2 edition, 2002.

[36] S. Ioffe, C. Szegedy, Batch normalization: accelerating deep network training by reducing internal covariate shift (2015).

[37] D. P. Kingma, J. Ba, Adam: a method for stochastic optimization (2014).

[38] A. Krogh, J. A. Hertz, A simple weight decay can improve generalization, M. I. Jordan, Y. LeCun, S. A. Solla (Eds.), Adv. Neural Inf. Process. Syst., volume 4, The MIT Press, 1992, pp. 950–957.

[39] TensorFlow v. r1.15, 2019.

[40] M. Tanahashi, M. Fujimura, T. Miyauchi, Coherent fine-scale eddies in turbulent premixed flames, Proc. Combust. Inst. 28 (2000) 529–535.

[41] Y. Nada, M. Tanahashi, T. Miyauchi, Effect of turbulence characteristics on local flame structure of H2-air premixed flames, J. Turbul. 5 (2004).

[42] Y. Shim, S. Tanaka, M. Tanahashi, T. Miyauchi, Local structure and fractal characteristics of H2–air turbulent premixed flame, Proc. Combust. Inst. 33 (2011) 1455–1462.

[43] Y. Shim, N. Fukushima, M. Shimura, Y. Nada, M. Tanahashi, T. Miyauchi, Radical fingering in turbulent premixed flame classified into thin reaction zones, Proc. Combust. Inst. 34 (2013) 1383–1391.

[44] Y. Minamoto, B. Yenerdag, M. Tanahashi, Morphology and structure of hydrogen–air turbulent premixed flames, Combust. Flame 192 (2018) 369–383.

[45] Y. Minamoto, N. Fukushima, M. Tanahashi, T. Miyauchi, T. D. Dunstan, N. Swaminathan, Effect of flow-geometry on turbulence-scalar interaction in premixed flames, Phys. Fluids 23 (2011) 125107.

[46] Y. Minamoto, K. Aoki, M. Tanahashi, N. Swaminathan, DNS of swirling hydrogen–air premixed flames, Int. J. Hydrogen Energy 40 (2015) 13604–13620.

[47] K. Aoki, M. Shimura, J. Park, Y. Minamoto, M. Tanahashi, Response of Heat Release Rate to Flame Straining in Swirling Hydrogen-Air Premixed Flames, Flow, Turbul. Combust. 104 (2020) 451–478.

[48] T. D. Dunstan, N. Swaminathan, K. N. C. Bray, N. G. Kingsbury, Flame Interactions in Turbulent Premixed Twin V-Flames, Combust. Sci. Technol. 185 (2013) 134–159.

[49] T. D. Dunstan, N. Swaminathan, K. N. C. Bray, Influence of flame geometry on turbulent premixed flame propagation: a DNS investigation, J. Fluid Mech. 709 (2012) 191–222.

[50] D. Endres, J. Schindelin, A new metric for probability distributions, IEEE Trans. Inf. Theory 49 (2003) 1858–1860.

[51] D. G. Goodwin, H. K. Moffat, I. Schoegl, R. L. Speth, B. W. Weber, Cantera: An Object-oriented Software Toolkit for Chemical Kinetics, Thermodynamics, and Transport Processes, 2017.

[52] G. P. Smith, D. M. Golden, M. Frenklach, N. W. Moriarty, B. Eiteneer, M. Goldenberg, C. T. Bowman, R. K. Hanson, S. Song, W. C. Gardiner, V. V. Jr., Lissianski, Z. Qin, GRI–Mech 3.0, 2000.

[53] A. Zhang, Z. C. Lipton, M. Li, A. J. Smola, Dive into Deep Learning, 2021.