

On the Use of Machine Learning for Subgrid Scale Filtered Density Function Modelling in Large Eddy Simulations of Combustion Systems



S. Iavarone, H. Yang, Z. Li, Z. X. Chen, and N. Swaminathan

Abstract The application of machine learning algorithms to model subgrid-scale filtered density functions (FDFs), required to estimate filtered reaction rates for Large Eddy Simulation (LES) of chemically reacting flows, is discussed in this chapter. Three test cases, i.e., a low-swirl premixed methane-air flame, a MILD combustion of methane-air mixtures, and a kerosene spray turbulent flame, are presented. The scalar statistics in these test cases may not be easily represented using the commonly used presumed shapes for modeling FDFs of mixture fraction and progress variable. Hence, the use of ML methods is explored. Particularly, deep neural network (DNN) to infer joint FDFs of mixture fraction and progress variable is reviewed here. The Direct Numerical Simulation (DNS) datasets employed to train the DNNs in each test case are described. The DNN performances are shown and compared to typical presumed probability density function (PDF) models. Finally, this chapter examines the advantages and caveats of the DNN-based approach.

S. Iavarone (✉)

Aero-Thermo-Mechanics Laboratory, Université Libre de Bruxelles, Brussels, Belgium
e-mail: si339@cam.ac.uk; salvatore.iavarone@ulb.be

S. Iavarone · H. Yang · Z. Li · N. Swaminathan
Engineering Department, University of Cambridge, Cambridge, UK
e-mail: hy345@cam.ac.uk

Z. Li
e-mail: zl443@cam.ac.uk

N. Swaminathan
e-mail: ns341@cam.ac.uk

Z. X. Chen
State Key Laboratory of Turbulence and Complex Systems, Aeronautics and Astronautics,
College of Engineering, Peking University, Beijing 100871, China
e-mail: chenzhi@pku.edu.cn; zc252@cam.ac.uk

Department of Engineering, University of Cambridge, Cambridge, UK

© The Author(s) 2023

N. Swaminathan and A. Parente (eds.), *Machine Learning and Its Application to Reacting Flows*, Lecture Notes in Energy 44, https://doi.org/10.1007/978-3-031-16248-0_8

1 Introduction

Increasingly stringent regulations on pollutants emissions from fossil fuel combustion are demanding for novel combustion technologies which can have high fuel flexibility, increased efficiency and low emissions. Moreover, a significant adoption of renewable technologies in future years is expected to reduce carbon footprint and meet the long-term objective of CO₂ neutrality. Nevertheless, combustion-based energy technologies will play a role in the future (or low-carbon) energy mix as discussed in the chapter “[Introduction](#)”. Hence, combustion research is called in to provide solutions to the expected challenges arising from issues related to fuel flexibility and improving efficiency with pollutants reduction. Current combustion studies focus on aspects such as development, validation and uncertainty quantification of new models, and involve either experiments or numerical simulations, or both. A collection of these studies represents a massive amount of data that can be leveraged to achieve significant progress in combustion science. Utilising this data has thus become a new challenge and research opportunity. Data-driven techniques such as machine learning (ML) have demonstrated their abilities to extract information from massive data and assist in developing novel models which can be leveraged for technology development.

Machine learning techniques allow us to have statistical inference, for some unknown quantities of interest, with reasonably accuracy and confidence by *carefully training* the algorithms using representative data. Since the 1990s, ML has regained increasing attention and achieved outstanding results in many areas (Jordan and Mitchell 2015), including science, technology, manufacturing, finance, education, health care, and many more. Combustion science is not an exception to this trend, there are many studies demonstrating successful use of ML for combustion and some of these studies date almost 30 years back. Christo and coworkers (Christo et al. 1995, 1996b, a) first employed a machine learning algorithm, namely the Artificial Neural Network (ANN), in the 1990s to deal with chemistry tabulation for turbulent combustion simulations. These works involved training an ANN to obtain changes in the composition of several reactive scalars rather than using the conventional direct integration of the relevant equations. Satisfactory results suggested that the ANN was able to provide, with computational efficiency, the chemical kinetics information required for turbulent combustion simulations. The computational efficiency was mainly noted to come from memory saving. The subsequent studies extended this novel approach to more complex chemical systems (Blasco et al. 1998, 1999; Chen et al. 2000), where multiple ANNs were proposed for different subdomains of the large composition space. The valuable time saving achieved by ANN compared with traditional methods was presented. The recent advances on ML applied to chemical kinetics are discussed in chapters “[Machine Learning Techniques in Reactive Atomistic Simulations](#)” and “[Machine Learning for Combustion Chemistry](#)” with different perspectives.

Blasco et al. (2000) employed two different ANNs, namely the Self-Organising Map (SOM) and the Multi-layer Perceptron (MLP), to estimate the thermochemical

states during a combustion simulation. The SOM was used to partition the thermochemical space into subdomains, while several MLPs were trained on each subdomain to predict the evolution of the thermochemical space in time. These early explorations identify a general route to utilise the ANN for chemistry tabulation approaches, although their generality was limited due to the similarity between training and testing cases. Consequently, later studies focused on developing ANNs for a wider range of combustion conditions.

Sen et al. trained ANNs using unsteady flame-turbulence-vortex interaction cases and subsequently used them for Large Eddy Simulations (LES) of syngas/air flames quite successfully (Sen and Menon 2009; Ali Sen and Menon 2010; Sen et al. 2010). Zhou et al. demonstrated successful application of the ANN to turbulent premixed flames by including 1D laminar premixed flame cases at different turbulent intensities while training the ANN (Zhou et al. 2013). A wider range of combustion conditions were also considered in later studies by including non-premixed laminar flamelets (Chatzopoulos and Rigopoulos 2013) to include local extinction and reignition (Franke et al. 2017) and non-adiabatic conditions (Wan et al. 2020, 2021) in the training data sets. Furthermore, randomising the non-premixed flamelets before using them as training data sets were shown to improve the generality of the ANN and helped to capture the behaviour of turbulent premixed flames quite well (Readshaw et al. 2021; Ding et al. 2021). Also, other techniques were explored to improve the generalisation level of ANN: Chi et al. (2021) trained the ANN on-the-fly during a simulation, whereas An et al. (2020) trained their ANN using data from Reynolds-averaged Navier–Stokes (RANS) simulations of hydrogen/carbon monoxide/kerosene/air mixture in a rocket combustion chamber and tested it for LES.

Further to the chemical kinetics use, another application of the ANN focuses on replacing the traditional flamelet look-up table, which requires a large memory. The general procedure is to set thermochemical scalars, which are the basis of the look-up table, as the input of the ANN and to infer the tabulated values. This reduces the memory requirement significantly since only the weights and bias(es) of the ANN need to be saved. A first successful application was demonstrated in Flemming et al. (2005) by building ANNs having the mixture fraction, its variance and its scalar dissipation rate as inputs and mass fractions as outputs, and using them in LES of the Sandia flame D. This was extended in Kempf et al. (2005) and Emami and Fard (2012) to estimate scalar mass fraction variations in a turbulent $\text{CH}_4/\text{H}_2/\text{N}_2$ jet diffusion flame. The optimisation of the ANN architecture, in terms of number of hidden layers and neurons per layer, was also explored to improve the predictive accuracy of LES of the Sydney bluff-body swirl-stabilised methane-hydrogen flame (Ihme et al. 2006, 2008, 2009).

The use of ANN for inferring multi-dimensional flamelet library is also explored in recent studies. Owoyele et al. proposed a grouped multi-targets ANN approach to model 4D and 5D flamelet libraries respectively for a *n*-dodecane spray flame, under conditions of the Spray A flame from the Engine Combustion Network (ECN), and methyl decanoate combustion in a compression ignition engine (Owoyele et al. 2020). Ranade et al. (2021) trained a SOM-MLP method on a 4D Probability Density

Function (PDF) table and used it for RANS and LES of the DLR-A turbulent jet diffusion flame. These works showed that the ANN yielded good accuracy at reduced computational costs with low storage space requirements. Similarly, Zhang et al. (2020) extended the application of the SOM-MLP algorithm to the Flamelet Generated Manifolds (FGM) model by using species mass fractions in mixture fraction-progress variable space as training data. This ANN approach was successfully used in RANS calculations and LES of ECN Spray H flame to explore the detailed spray combustion process. More comprehensive reviews of the applications of ML in combustion research can be found in Zheng et al. (2020), Zhou et al. (2022) and Ihme et al. (2022).

Presumed PDF shapes are typically used along with tabulated chemistry approaches. The PDF of relevant scalars such as mixture fraction and progress variable are used to compute averaged temperature, density, species mass fractions, and the relevant reaction rates. These quantities can be stored in a look-up table with the first two moments of the above scalars as controlling variables. Although widely employed in several past studies, presumed PDF or Filtered Density Function (PDF), in the context of LES, approaches may not accurately represent the scalar statistical behaviour under several conditions, such as extinction and reignition, combustion among multiple streams, multi-regime burners, and multi-phase reacting flows. The PDFs having shapes different to the regular distributions such as Gaussian or β -function can be also observed prominently in Moderate or Intense Low-oxygen Dilution (MILD) combustion. This combustion mode features broadly distributed reaction zones rather than conventional flamelet-like structures, with strong interactions between autoigniting and propagating fronts. Therefore, it may not be satisfactory to use conventional PDFs/PDFs models to predict reaction rates, and advanced data-driven techniques like machine learning may be a suitable alternative for improving the accuracy. De Frahan et al. (2019) compared the performance of three different machine learning techniques, *viz.*, random forests, which is a traditional ensemble methods, deep neural networks (DNNs), and conditional variational autoencoder (CVAE), multiple hidden layers between which is also known as generative learning, to infer marginal PDFs of reaction progress variable in a swirling methane/air premixed flame and showed that DNN is superior compared to the other two techniques. The DNN is an ANN with multiple hidden layers between input and output. Yao et al. (2020) built an MLP to obtain the mixture fraction marginal PDF for LES of turbulent spray flames and observed an order of magnitude improvement compared to those of the traditional presumed PDF approaches. Chen et al. (2021) employed a DNN to predict the joint PDF of mixture fraction and progress variable in MILD combustion conditions and showed that the DNN is generally able to capture the complex PDF behaviours and their variations with excellent accuracy, outperforming other presumed PDF models.

This chapter aims to provide an overview of recent studies employing deep neural networks (interchangeably referred to as DNN, ANN or MLP hereafter) to infer subgrid-scale PDFs and reaction rates needed for LES of turbulent combustion under conventional and MILD conditions. A review of the Direct Numerical Simulation (DNS) data used to train these DNNs is also given. The chapter is structured as

follows. A recap of the treatment of FDFs in LES of turbulent combustion systems is provided in Sect. 2. The DNS cases used as training datasets for the DNNs are described in Sect. 3. The characteristics of the DNNs employed for the different combustion cases are illustrated in Sect. 4. The main results in terms of FDF and reaction rate predictions are discussion in Sect. 5. The conclusions are summarised in Sect. 6.

2 FDF Modelling

The filtered reaction rate appearing in the transport equation for a species filtered mass fraction or reaction progress variable needs a closure model and recent developments in various closure models are described in the book (Swaminathan et al. 2022) and review papers (Veynante and Vervisch 2002; Pitsch 2006). Earlier chapters of this book discuss the potential application of ML techniques to some of the reaction rate closures. In the presumed PDF approach, the filtered reaction rate is modelled as an integral of the product of a conditional reaction rate and a FDF (see Eq. 6). The mixture fraction and the reaction progress variable are typically used as conditioning variables to signify the role of mixing and flame propagation on reaction rate (Bradley et al. 1998; Ihme and Pitsch 2008a). The conditional reaction rate may be estimated using one of the methods developed in past studies and these methods used canonical flames for chemistry tabulation, e.g., flamelet-generated manifolds (van Oijen and de Goey 2002), flame prolongation of intrinsic low dimensional manifold (Gicquel et al. 2000), conditional source term estimation method (Jin et al. 2008), or the solution of conditionally filtered equations for species mass fractions and energy via the conditional moment closure method (Klimenko and Bilger 1999).

The subgrid variations in the conditioning variables about their filtered values are represented by the filtered density function (FDF). The FDF can generally be obtained by solving its transport equations using various approaches, e.g., Lagrangian particles (Pope 1985), Eulerian stochastic fields (Jones and Kakhi 1998), and multi-environment (Fox 2003). However, these approaches are computationally expensive and thus using a presumed FDF can be chosen (Pitsch 2006; Pope 2013) to save computational costs. This presumed FDF approach will need only the statistical moments, usually the mean and variance, of the key variables (mixture fraction, progress variable, flame stretch/straining, heat loss, etc., depending on the physical scenario of interest) to be transported and it is therefore much more economical.

The β -PDF (Cook and Riley 1994) is the most commonly used presumed FDF in LES of turbulent flames (Raman et al. 2005; Navarro-Martinez et al. 2005; Ihme and Pitsch 2008b; Chen et al. 2017), and it usually provides a good approximation of a conserved scalar distribution. The Favre-averaged FDF of the mixture fraction Z with a presumed β -distribution is calculated as

$$\tilde{\rho}_\beta(\xi; \tilde{Z}, \tilde{\sigma}_Z^2) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \xi^{a-1} (1-\xi)^{b-1}, \quad (1)$$

where ξ is the sample space variable for Z , \tilde{Z} is the filtered mixture fraction and $\tilde{\sigma}_Z^2 \equiv \tilde{Z}'' = (Z - \tilde{Z})^2$ is the mixture fraction subgrid variance. The parameters of the Γ function are $a = \tilde{Z} (1/\tilde{g}_Z - 1)$ and $b = (1 - \tilde{Z}) (1/\tilde{g}_Z - 1)$. The segregation factor is $\tilde{g}_Z = \tilde{\sigma}_Z^2 / (\tilde{Z}(1 - \tilde{Z}))$. The Favre-filtered FDF of the progress variable, $\tilde{P}_\beta(\eta; \tilde{c}, \tilde{\sigma}_c^2)$, can also be presumed to follow a β distribution and obtained in a similar manner using \tilde{c} and $\tilde{\sigma}_c^2 \equiv \tilde{c}'' = (c - \tilde{c})^2$. The joint FDF of ξ and η can be modelled as

$$\tilde{P}(\xi, \eta) = \tilde{P}_\beta(\xi; \tilde{Z}, \tilde{\sigma}_Z^2) \tilde{P}_\beta(\eta; \tilde{c}, \tilde{\sigma}_c^2), \quad (2)$$

assuming that there is a weak correlation between the subgrid fluctuations of Z and c . Such assumption has been widely accepted for LES of conventional combustion (Pitsch 2006; Veynante and Vervisch 2002). However, stronger subgrid correlations of scalars fluctuations can occur in MILD combustion (Minamoto et al. 2014) and hence the above assumption may not applicable universally. Other analytical distributions have been considered in past studies (Grout et al. 2009; Darbyshire and Swaminathan 2012; Linse et al. 2014). Darbyshire and Swaminathan (2012) proposed a correlated joint PDF model using the *Plackett copula* (Plackett 1965) to include the covariance of Z and c in RANS calculations. The covariance, $\tilde{\sigma}_{Zc}$, written as $\tilde{\sigma}_{Zc} = \widetilde{(Z - \tilde{Z})(c - \tilde{c})}$ is used in the *copula* method to obtain a joint PDF from the univariate marginal distributions, $\tilde{P}_\beta(Z)$ and $\tilde{P}_\beta(c)$. For non-zero values of $\tilde{\sigma}_{Zc}$, the correlated joint PDF is calculated as

$$\tilde{P}(Z, c) = \frac{\theta \tilde{P}_\beta(Z) \tilde{P}_\beta(c) (\mathcal{A} - 2\mathcal{B})}{(\mathcal{A}^2 - 4\theta\mathcal{B})^{3/2}}, \quad (3)$$

with

$$\mathcal{A} = 1 + (\theta - 1) [\tilde{\mathcal{E}}_\beta(Z) + \tilde{\mathcal{E}}_\beta(c)], \quad (4)$$

and

$$\mathcal{B} = (\theta - 1) \tilde{\mathcal{E}}_\beta(Z) \tilde{\mathcal{E}}_\beta(c), \quad (5)$$

where $\tilde{\mathcal{E}}_\beta$ is the β cumulative distribution function (CDF) and θ is the odds ratio calculated using a Monte Carlo approach (Ruan et al. 2014). The *copula* method has been used in RANS calculations of stratified premixed and lifted jet flames (Ruan et al. 2014; Chen et al. 2015) showing improved prediction of the lift-off height with respect to the double- β PDF given in Eq. (2).

In presumed-FDF approaches, the subgrid reaction rate is obtained as

$$\bar{\omega} = \int_0^1 \int_0^1 \langle \dot{\omega} | Z, c \rangle P(Z, c; \tilde{Z}, \tilde{\sigma}_Z^2, \tilde{c}, \tilde{\sigma}_c^2) dZ dc, \quad (6)$$

and this approach reduces the computational cost significantly for LES by using presumed FDF in the above equation. However, the presumed FDF shapes obtained using

classical functions, for example bimodal delta function, may not be fully satisfactory for situations such as (i) MILD combustion conditions, (ii) when there are evaporating droplets, and (iii) when the burnt or burning mixture is inhomogeneous leading to significant statistical correlation between Z and c (Chen et al. 2018). To overcome these issues, machine learning algorithms are employed to construct predictive models for the scalar PDFs/FDFs in recent studies. A deep neural network (DNN), among other ML techniques tested, was shown to be better than a joint β -function model in inferring subgrid FDFs in a swirling methane-air premixed flame (de Frahan et al. 2019). This behaviour was also demonstrated for MILD combustion (Chen et al. 2021) and turbulent spray flames (Yao et al. 2020). These tests were conducted using respective direct numerical simulation (DNS) datasets. DNS can be seen as a *virtual* experiment resolving all the relevant length and time scales without turbulence modelling. Thus, it is a powerful tool for investigating combustion models. It is quite straightforward to obtain filtered quantities from DNS data by applying appropriate filtering operations (Pope 2000) and these can be used as input to ML algorithms such as DNN. The data extraction and its processing prior to using them for DNN training are important steps which can play a role to improve accuracy and generality of the neural networks. Details about these steps, along with the main features of the cases studied in de Frahan et al. (2019), Chen et al. (2021) and Yao et al. (2020), are discussed in the following sections. Details on the respective DNS cases can be found in those studies as the objective here is on the use of ML techniques.

3 DNS Data Extraction and Manipulation

Three combustion cases are considered in this chapter: a low-swirl premixed methane-air flame investigated in de Frahan et al. (2019), methane-air combustion under MILD conditions studied in Chen et al. (2021), and a turbulent kerosene spray flame used in Yao et al. (2020). The corresponding DNS setups and data preparation procedures are described next.

3.1 Low-Swirl Premixed Flame

The DNS dataset considered by de Frahan et al. is a snapshot of a quasi-stationary simulation of an experimental low-swirl, premixed methane-air burner (Day et al. 2012). In this setup, a nozzle imposes a low swirl to a CH_4 /air mixture with fuel-air equivalence ratio $\phi = 0.7$ at the inflow. The nozzle region is surrounded by a co-flow of cold air. A lifted premixed flame with its partially burnt mixture reacting with co-flow air in downstream locations was observed in the experiments. The presence of this multi-regime burning introduces challenges for modeling the joint FDF of mixture fraction and progress variable. Training ML models with such DNS dataset has additional advantages such as using diverse subsets as training data, avoiding

overfitting, and increasing the opportunities for model generalisation. The training sets were constructed by selecting different subvolumes, indicated by \mathcal{V} as in Fig. 1, spanning from premixed combustion region to downstream zone containing mixing of premixed combustion products with co-flow air. de Frahan et al. (2019) used a single time snapshot at $t = 0.0626$ s from the DNS to demonstrate the capabilities of ML for FDF modelling. In the context of LES, the FDF at a given point and time can be extracted by applying fine-grained filtering to DNS or experimental data at a given instant (Pope 1990). In each subvolume, sample moments and the associated FDF were thus obtained by using a discrete box filter:

$$\overline{\psi}(x, y, z) = \frac{1}{n_f^3} \sum_{i=-n_f/2}^{n_f/2} \sum_{j=-n_f/2}^{n_f/2} \sum_{k=-n_f/2}^{n_f/2} \psi(x + i\Delta x, y + j\Delta x, z + k\Delta x), \quad (7)$$

where ψ is the quantity of interest, n_f is the number of points in the discrete box filter, $\overline{\Delta} = 32\Delta x$ is the filter size, and $\Delta x = 100 \mu\text{m}$ is the smallest spatial cell size in the DNS (six times smaller than the laminar flame thickness). Four sample moments of the joint FDF, i.e., \tilde{Z} , $\tilde{\sigma}_Z^2$, \tilde{c} , $\tilde{\sigma}_c^2$, which are Favre-filtered mixture fraction, its subgrid scale (SGS) variance, progress variable and its SGS variance, were extracted for each subvolume. The filter size was chosen to be representative of typical LES filter scale (Pitsch 2006) and to ensure adequate samples to construct FDF. These filters were spaced equidistant of $8\Delta x$, leading to 58800 FDFs for each subvolume. The mixture fraction Z was defined using nitrogen mass fraction so that it took a value of 1 in the burner stream and 0 in the co-flow air. The progress variable, varying between 0 and 0.21, was defined using mass fractions of CO_2 , CO , H_2O and H_2 as $c = Y_{\text{CO}_2} + Y_{\text{CO}} + Y_{\text{H}_2\text{O}} + Y_{\text{H}_2}$. The density-weighted FDFs of Z and c were constructed using 64 bins in Z space and 32 bins in c space, which gives a vector of 2048 values to describe a single joint FDF. The conditional means of the reaction rate $\langle \dot{\omega} | Z, c \rangle$ were also extracted for each sample with an identical discretisation. Prior to training, the sample moments were independently centered by subtracting the median and scaled by dividing the data by the range between the 25th and 75th quantiles. It is known that appropriate centring and scaling are generally beneficial for ML algorithms (Goodfellow et al. 2016). According to the authors this centring and scaling were robust to outliers. The samples from a volume \mathcal{V}_i were randomly split among two distinct datasets: a training dataset, \mathcal{D}_i^t , and a validation dataset, \mathcal{D}_i^v , comprising 5% of the total samples, as illustrated in Fig. 1.

3.2 MILD Combustion

The MILD combustion DNS dataset of Doan et al. (2018) was used to study the application of DNN for inferring subgrid FDF in MILD combustion by Chen et al. (2021). A cube of size $L_x \times L_y \times L_z = 10 \times 10 \times 10$ mm was used to conduct DNS of turbulent combustion of inhomogeneous methane-air mixtures diluted with exhaust

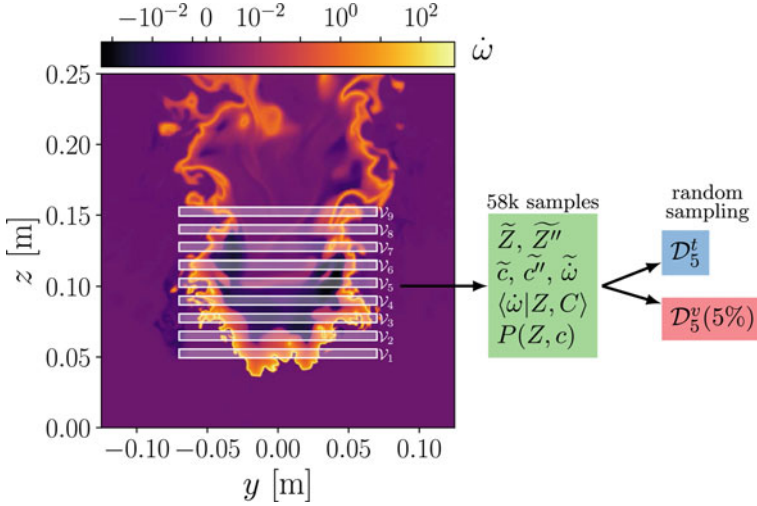


Fig. 1 Illustration of data generation procedure for \mathcal{V}_5

gases. A spatial resolution of $\delta x \approx 20 \mu\text{m}$ obtained using 512 points distributed uniformly in each direction was observed to be sufficient to resolve the turbulent and chemical length scales of interest as described in Doan et al. (2018). The simulation was run for 1.5 flow-through time τ_f , defined in Minamoto and Swaminathan (2015). Further detail on the DNS procedure and datasets can be found in Doan et al. (2018). Three cases, *viz.*, AZ1, AZ2 and BZ1, with different mixing length scales and dilution levels were considered for the DNN training. The conditioning variables for the FDF analyses were the Bilger mixture fraction (Bilger 1976) and a temperature-based reaction progress variable, c_T , defined as

$$c_T = \frac{T - T_u}{T_b(Z) - T_u}, \quad (8)$$

where T_u is 1500 K and the value of burnt mixture temperature T_b depends on Z and it can be obtained using MILD Flame Element (MIFE) laminar calculations (Minamoto and Swaminathan 2014). Favre-filtered fields were extracted from the DNS by applying a low-pass box filter. For example, the Favre-filtered mixture fraction \tilde{Z} was obtained as:

$$\tilde{Z}(\mathbf{x}, t) = \frac{1}{\bar{\rho}(\mathbf{x}, t)} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \rho(\mathbf{x}', t) Z(\mathbf{x}', t) d\mathbf{x}', \quad (9)$$

where $\bar{\cdot}$ and $\tilde{\cdot}$ denote the Reynolds and Favre filtering respectively, ρ is the mixture density and Δ is the filter width. The position vectors are \mathbf{x} and \mathbf{x}' . The subgrid variance was obtained as

$$\widetilde{\sigma_Z^2}(\mathbf{x}, t) = \frac{1}{\overline{\rho}(\mathbf{x}, t)} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \rho(\mathbf{x}', t) [Z(\mathbf{x}', t) - \widetilde{Z}(\mathbf{x}, t)]^2 d\mathbf{x}'. \quad (10)$$

Similarly, the \widetilde{c}_T and $\widetilde{\sigma_{c_T}^2}$ fields were calculated as above. The Z - c_T joint FDF was then computed as

$$\widetilde{P}(\xi, \eta; \mathbf{x}, t) = \frac{1}{\overline{\rho}(\mathbf{x}, t)} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \rho(\mathbf{x}', t) \delta[\xi - Z(\mathbf{x}', t)] \delta[\eta - c_T(\mathbf{x}', t)] d\mathbf{x}', \quad (11)$$

where ξ and η were the sample-space variables of Z and c_T respectively, $\delta[\cdot]$ is the Dirac delta function. The discrete FDFs were obtained for a given point in a given DNS snapshot by binning the Z and c_T samples in the corresponding filtering subspace with 35 non-uniform bins in Z space (clustered around the stoichiometric value) and 31 uniform bins in c_T space. The subgrid-scale covariance, $\widetilde{\sigma_{Zc_T}}$, also used by the *copula* model, was computed as

$$\widetilde{\sigma_{Zc_T}}(\mathbf{x}, t) = \frac{1}{\overline{\rho}(\mathbf{x}, t)} \int_{x-\frac{\Delta}{2}}^{x+\frac{\Delta}{2}} \rho(\mathbf{x}', t) [Z(\mathbf{x}', t) - \widetilde{Z}(\mathbf{x}, t)] \times [c_T(\mathbf{x}', t) - \widetilde{c}_T(\mathbf{x}, t)] d\mathbf{x}'. \quad (12)$$

The filtered scalar fields \widetilde{Z} , \widetilde{c}_T , $\widetilde{\sigma_Z^2}$, $\widetilde{\sigma_{c_T}^2}$ and $\widetilde{\sigma_{Zc_T}}$ formed the DNN input matrix \mathbf{X} . The unfiltered ρ , Z and c_T fields were used to obtain the Favre filtered FDFs required for the target matrix \mathbf{Y} . The procedure is shown schematically in Fig. 2 for a snapshot of case AZ1. The filtered fields are presented in 2D with the thin DNS grid-lines for visual clarity. The indices i , j and k pertain to the x , y and z directions in 3D space, respectively, and are assigned to each ‘‘LES filter cube’’ indicated by a red box in Fig. 2. The total number of samples taken in each direction is n_{cube} . The effects of filter size were also investigated by considering a range of filter sizes relevant to typical LES. The filter sizes were normalized using the thermal thickness of the stoichiometric MIFE, $\delta_{\text{th}}^{\text{st}} = 1.6$ mm. A filter size of $\Delta = 80\delta x$ corresponded to $\Delta^+ = \Delta/\delta_{\text{th}}^{\text{st}} = 1$. The extracted matrices \mathbf{X} and \mathbf{Y} were flattened to be two-dimensional, with as many rows as the number of samples and as many columns as the number of features. The input matrix \mathbf{X} had 5 columns, while the target matrix \mathbf{Y} had 1085 columns, obtained from the discretisation step mentioned above.

Centring and scaling of the input matrix \mathbf{X} were performed as follows: each column vector, having n_{cube}^3 elements, was centred by subtracting its mean and scaled by dividing by its standard deviation. Centring and scaling were not applied to the output matrix \mathbf{Y} . However, to address the issue of having unbounded values of the FDFs, the discrete density function values were considered. As such, every number in \mathbf{Y} varied between 0 and 1, and the sum of the elements in each target row is equal to 1.

Subsequent to the scaling procedures, a dimensionality reduction technique like Principal Component Analysis (PCA), discussed in chapter ‘‘[Reduced-Order Mod-](#)

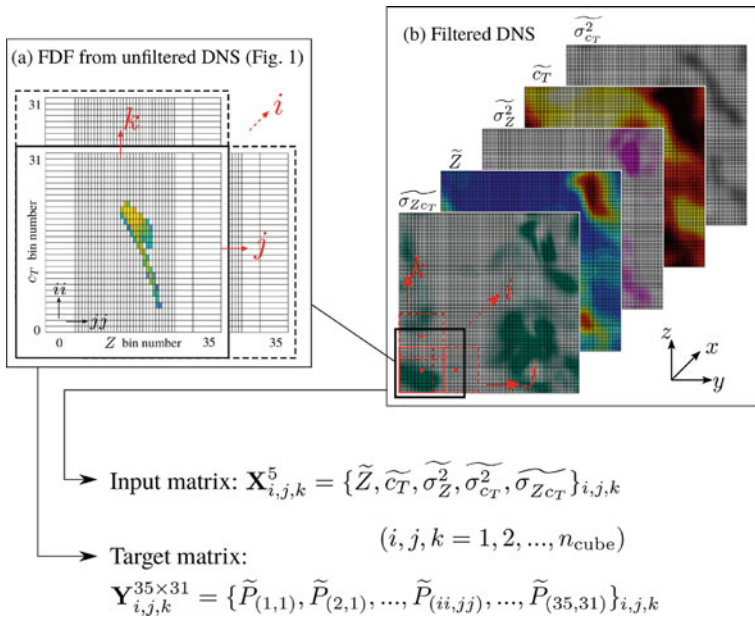


Fig. 2 Schematic demonstration of the construction of the DNN input and target matrices (Chen et al. 2021)

eling of Reacting Flows Using Data-Driven Approaches” was used to identify and remove the outliers in the training data. Two types of outliers, *viz.*, *leverage* and *orthogonal*, Verdonck et al. (2009) were determined and discarded. Details about the identification and removal step are provided in Chen et al. (2021). Once leverage and orthogonal outliers were removed from the dataset, the DNN training was then performed on the remaining observations as discussed in the following Sect. 4.2.

3.3 Spray Combustion

Carrier-phase DNS (CP-DNS) data of turbulent spray flames were used to build a deep learning training database for mixture fraction FDF predictions. In carrier-phase DNS, the flow field is resolved with a point source approximation for the droplets, thus all relevant scales of the fluid phase are resolved except the boundary layers around individual particles. The governing equations of the gas phase are solved in the Eulerian framework and coupled with a Lagrangian solver for displacement, size, and temperature of the droplets. An equilibrium state of the liquid and the vapor at the interface was assumed. A full description of the governing equations is provided in Yao et al. (2020). The computational domain is a rectangular box, discretised by a mesh with $192 \times 128 \times 128$ cells having $\delta_{\text{DNS}} = 100 \mu\text{m}$. This grid size ensured a suffi-

cient resolution of the small scale structures of the flow field (Pope 2000), whereas a finer resolution could compromise the point particle assumptions of the liquid phase. Kerosene droplets (treated as single-component $C_{12}H_{23}$) were randomly injected into humid air, representative of experimental (Khan et al. 2007; Wang et al. 2018) and numerical (Wright et al. 2005; Giusti et al. 2018) setups. A homogeneous isotropic turbulent velocity field, calculated by a modified von Karman spectrum (Wang et al. 2019) was imposed at the inlet. The progressive kerosene droplet evaporation led to an ignitable mixture that promoted a statistically planar turbulent partially premixed flame. Further downstream, the hot post-flame temperatures led to reduced turbulence levels due to higher viscosity and a sudden evaporation of remaining droplets that could penetrate the flame. This lack of homogeneity and the presence of a source term for the mixture fraction are prone to make the existing FDF models (O'Brien and Jiang 1991; Cook and Riley 1994) inaccurate.

Filter boxes were used for post-processing of CP-DNS data to group several DNS cells into one LES cell. A filter box example is shown in Fig. 3 along with the DNS domain and setup, and the simulated temperature contour. The mixture fraction FDF $P(\eta)$ was computed from DNS data using a mixture fraction binning, with a bin size of 0.01 for all DNS cells lying within a specific LES cell. Favre filtering was used to extract LES quantities that were employed as input variables for the ANN. According to Klimenko and Bilger (1999), the following input quantities were found to have an effect on the mixing statistics and were thus considered: mixture fraction ξ , eddy viscosity ν_t , turbulence dissipation rate ϵ_t , diffusion coefficient D , density ρ , spray evaporation rate J_m , relative velocity between the droplet and the

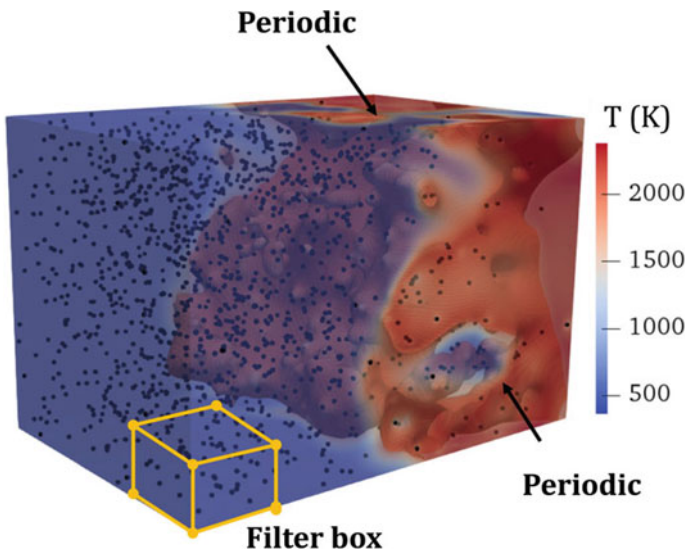


Fig. 3 Simulation setup of CP-DNS (solid points: droplets; the gas phase is colored by temperature) and an LES filter box (Yao et al. 2020)

surrounding gas U_d and droplet number density C . The turbulence dissipation rate was replaced by the more easily available strain rate $|S_{ij}|$. All the DNN inputs were filtered and Favre averaged. Therefore, the input features are commonly accessible in a typical LES of spray combustion. Moreover, Wang et al. concluded in their study that these parameters sufficiently characterize the mixture fraction FDF in turbulent spray flames. To ensure the reliability of the DNN for a reasonable range of LES meshes, the authors investigated the following LES filter sizes: $(\Delta_{LES})^3 = (8\delta_{DNS})^3$, $(\Delta_{LES})^3 = (16\delta_{DNS})^3$, $(\Delta_{LES})^3 = (32\delta_{DNS})^3$. The final database is a combination of data samples with different Δ_{LES} . The performance of the DNN for data samples using different LES filter boxes were assessed. The output target was set to be a placeholder of 60 elements covering ξ in $[0, 0.6]$, as $\xi_{max} \leq 0.6$ in the spray flame simulations. To avoid that the binning procedure could lead to empty bins, especially for small Δ_{LES} , missing values were replaced by interpolated values computed by Stineman interpolation method, which is widely used in statistics to deal with the missing values as it preserves the monotonicity of data and prevents introducing spurious oscillations (Stineman 1980). It was found that the commonly used zero-padding operation, which fills in blank data with zeros, is not applicable as the DNN would be misled and learn erroneous patterns. A total of 18 simulation cases were run to form the full database for training and validation purposes. The validation (test) dataset consisted of five simulation cases, resulting in a test/train ratio of about 0.38. These datasets included parameter ranges that approximate conditions to be expected in real spray flames and were used for the a priori validation presented in Sect. 5.

To recap, the three studies selected several DNS cases to construct a heterogeneous training set. If only one DNS case was available then several subdomains within the DNS domain were selected. Chen et al. (2021) considered one additional DNN input feature, i.e., the scalar covariance, to the input set chosen by de Frahan et al. (2019). Yao et al. (2020) chose different DNN input features specifically for spray combustion. No scaling was adopted by Yao et al., whereas two different scaling methods were implemented in the other studies. Only Chen et al. adopted an outlier removal by using a dimensional reduction technique. Discrete density functions, bounded between 0 and 1, were the DNN target in de Frahan et al. (2019) and Chen et al. (2021) while Yao et al. (2020) considered probability density function values. The review of these studies shows that no unique algorithm needs to be adopted to prepare the input data for a ML model. The only common goal that needs to be pursued is to construct an input dataset that is as heterogeneous as possible to increase the generalisation, also known as transfer learning, of the trained ML models. The similarities and differences of the DNNs used in these three studies are discussed next.

4 Deep Neural Networks for Subgrid-Scale FDFs

A standard neural network consists of many simple connected functional units, called neurons. Each neuron receives an input which is processed through activation functions to produce an output. Multiple neurons can be combined to form fully connected networks, which are called artificial neural networks (ANNs) since they mimic the neuron arrangements in the human brain. Feed-forward networks, also called multi-layer perceptrons (MLPs), are classic ANN structures, and they are composed of layers of neurons, where a weighted output from one layer is the input to the next layer. The first layer of the MLP accepts a vector as input and the elements of this vector are known as features. The final output of the MLP is the target quantity of interest. The layer providing the final MLP output is called output layer, while the other layers in the network are called hidden layers. In a mathematical perspective (Goodfellow et al. 2016), the MLP defines a mapping from the input \mathbf{x} to the output $\mathbf{y} = f(\mathbf{x}, \boldsymbol{\theta})$, where the parameters $\boldsymbol{\theta}$ are the trainable network parameters. Each neuron is a functional unit that is generally described by

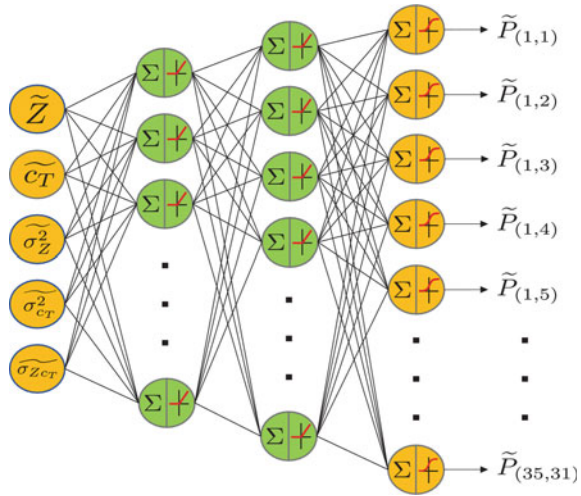
$$\mathbf{y} = \phi(\mathbf{x}^T \boldsymbol{\omega} + \mathbf{b}), \quad (13)$$

where $\boldsymbol{\omega}$ and \mathbf{b} are the weights and bias vector, and ϕ is the activation function (see Sect. 2.3.7.2, Chap. 2, this volume), which provides great flexibility to ANNs by introducing non-linearity to an otherwise linear relationship between input and output. There are several activation functions and some of these will be introduced and described later. The weight $\boldsymbol{\omega}$ is a matrix of the size $k \times m$, whereas the bias \mathbf{b} is a vector of m elements. For each layer, k is the number of inputs received from the preceding layer and m is the number of neurons in the current layer. $\boldsymbol{\omega}$ and \mathbf{b} contains the trainable parameters of the network. The training of ANNs pursues the objective of minimizing a target loss function

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\omega}) = \mathcal{G}(f(\mathbf{x}, \boldsymbol{\omega}) - f^*), \quad (14)$$

where \mathcal{G} is any measure of the difference between the modeled value f and the real value f^* . The most commonly used loss functions are the mean absolute error (MAE) and the mean squared error (MSE). Nonlinear optimization methods, such as backward propagation (Rumelhart et al. 1986), are used to identify the network weights that minimize the error between predictions and labeled training data. The training step gives the optimized set of weights. The MLP is a design that is suitable for regression problems, whereas other types of ANNs, such as Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), have been extensively used in processing image data and time series problems, etc., see Sect. 2.3.7.2 (Chap. 2, this volume), for further detail. A schematic of the MLP architecture with input, hidden, and output layers is shown in Fig. 4 as an example.

Fig. 4 A schematic of 3-layer MLP architecture



4.1 Low-Swirl Premixed Flame

A feed-forward fully connected DNN with three, two hidden and an output, layers was trained by de Frahan et al. (2019) to predict the joint subfilter FDF of mixture fraction and progress variable. There were 256 and 512 neurons in the two hidden layers and neurons had a leaky rectified linear unit activation function (LeakyReLU):

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0 \\ \alpha x_i & \text{otherwise} \end{cases} \tag{15}$$

where x_i is the weighted sum of the neuron input, y_i is its output, and α , usually equal to 0.01, is the slope. A LeakyReLU activation function avoids mapping negative input to zero values unlike its parent function ReLU having $\alpha = 0$. A large weight update during training can yield the summed input to be always negative regardless of the network input. A neuron featuring a ReLU function will then output a zero value leading to the *dying* ReLU case, in which the neuron neither activates a gradient-based optimization nor adjust its weights. Furthermore, similar to the vanishing gradients problem, the learning can be slow while training ReLU networks stumbling on constant zero gradients. The leaky rectifier allows for a small, non-zero gradient when the unit is saturated and not active. Additionally, each hidden layer is followed by a batch normalization layer (Ioffe and Szegedy 2015) and this technique has been widely used to build deep networks as it leads to speed and performance improvements. It applies the following function:

$$y_i = \gamma \frac{x_i - \mu_x}{\sqrt{\sigma_x^2 + \epsilon}} + \delta \tag{16}$$

where x_i and y_i are the i -th elements of the layer input and output vectors respectively. These vectors are of size n having a mean and variance of $\mu_x = 1/n \sum_{i=1}^n x_i$ and $\sigma_x^2 = 1/n \sum_{i=1}^n (x_i - \mu_x)^2$. A small real number ϵ is used to maintain numerical stability. Both γ and δ are learning parameter vectors of size n and they are updated iteratively during training for optimization purposes. de Frahan et al. (2019) chose $\epsilon = 10^{-5}$ and a moving average of μ_x and σ_x computed during training with a decay of 0.1 (or, equivalently, momentum of 0.9).

The DNN inputs are the four moments of the joint FDF, viz., \tilde{Z} , $\tilde{\sigma}_Z^2$, \tilde{c} , and $\tilde{\sigma}_c^2$ whereas the outputs are a total of 2048 FDF values obtained from the discretisation of the joint FDF of mixture fraction Z and progress variable c as described in Sect. 3.1. Thus, an output layer having 2048 neurons, as many as the number of outputs, was considered in de Frahan et al. (2019). The output layer features a *softmax* activation function:

$$y_i = \frac{\exp(x_i)}{\sum_{i=1}^n \exp(x_i)} \quad (17)$$

where x_i and y_i are defined as for Eq. 16. This type of activation function ensures that $\sum_{i=1}^n y_i = 1$ and $y_i \in [0, 1] \forall i$. The loss function used was the binary cross entropy between the target y and the prediction \hat{y} and this function is

$$\mathcal{L}(\hat{y}, y) = \frac{1}{n} \sum_{i=1}^n (y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i)), \quad (18)$$

representing a proper metric for measuring the difference between two probability distributions. The total number of trainable parameters was 1.1 M. The training was performed over 500 epochs, i.e., 500 training loops through the entire training data. For each epoch, the training data is fully shuffled and divided into batches with 64 training samples per batch. All trainable parameters are updated after each epoch. A split of 95/5% between training and validation samples was applied on the entire dataset. The loss function is computed on the validation samples which are not part of the training process. Thus, the validation loss is the true indicator of the ANN's performance and provides hints regarding its generality. It is a common practice to track the losses during both training and validation steps continuously to check if the losses are decreasing over each epoch by studying learning curves (a plot of loss vs epoch number). These learning curves can be used to diagnose an underfit, overfit, or well-fit model and whether the training or validation datasets are not representative of the problem domain. A good ANN training gives loss curves that decreases continuously until a plateau is reached where the difference between the training and validation losses is small. de Frahan et al. (2019) chose Adam optimizer (Kingma and Ba 2014), which is a gradient descent algorithm, with an initial learning rate of 10^{-4} . The learning rate is a dimensionless parameter that determines the step size of the stochastic gradient descent used to adjust the weights, ω . The Adam optimizer is more sophisticated than traditional stochastic gradient

descent by having a per-parameter learning rate, which can also be adapted during the training (Kingma and Ba 2014).

4.2 MILD Combustion

Chen et al. (2021) used a feed-forward fully connected DNN to infer the joint FDF of mixture fraction and progress variable. This DNN is similar to the one employed by de Frahan et al. (2019) and can be summarized as follows:

- linear hidden layer with 5 input features and bias, LeakyReLU activation function with $\alpha = 0.01$, and 256 output features;
- batch normalization layer with 256 input and output features, and momentum equal to 0.9;
- linear hidden layer with 256 input features and bias, LeakyReLU activation function with $\alpha = 0.01$, and 512 output features;
- batch normalization layer with 512 input and output features, and momentum equal to 0.9;
- linear output layer with 512 input features and bias, *softmax* activation function, and 1085 output features.

Thus, the two hidden layers had 256 and 512 fully connected neurons, where LeakyReLU activation functions were applied. Each hidden layer was followed by a batch normalization layer. The output layer contained 1085 neurons featuring a *softmax* activation function. The loss function used was the binary cross entropy given in Eq. 18 along with Adam optimizer with an initial learning rate of 10^{-4} . The model was trained for maximum 1000 epochs with batch size of 256 training samples. The ANN features were the four moments of the joint FDF and the outputs were a total of 1085 FDF values. A split of 80/20% between training and validation samples was applied on the entire dataset containing about 28000 filtered DNS boxes. An early stopping method, by using a predefined number of epochs, was used for the training to avoid overfitting. An overfitted ANN will have a validation loss that decreases for the first several epochs but increases subsequently (Goodfellow et al. 2016).

4.3 Spray Flame

Yao et al. (2020) used an MLP with four hidden layers and 500 neurons per layer to infer the Favre-filtered FDF of the mixture fraction in spray flames. As noted in Sect. 3.3, the input quantities were $\tilde{\xi}$, \tilde{v}_t , $|\widetilde{S_{ij}}|$, \tilde{D} , $\tilde{\rho}$, spray evaporation rate \tilde{J}_m , relative velocity between the droplet and the surrounding gas \tilde{U}_d , and droplet number density \tilde{C} . The output was a vector with 60 elements since the FDF of the mixture fraction $P(\eta)$ (where η is the sample space variable for the mixture fraction ξ) was obtained as described in Sect. 3.3. The activation function $\phi(z) = \max(0, z)$ applied

in each layer was the ReLU. A traditional stochastic gradient descent algorithm was used to minimize the mean absolute error, which was the loss function. A total of 18 DNS cases were run to form the full datasets for the training and validation steps. The validation (test) dataset consisted of five cases, resulting in a test/train ratio of ~ 0.38 . An early stopping criterion was imposed for the training process. This ANN was also trained on the conditional scalar dissipation rate $\langle N | \xi = \eta \rangle$, which is another interesting application.

5 Main Results

5.1 FDF Predictions and Generalisation

An overview of the ML model performance in each of the test cases is discussed in this section. The FDF predictions provided by ML and analytical models were assessed *a priori* using the FDFs obtained from the DNS cases.

5.1.1 Premixed Flame

Three different ML models, i.e., random forest (RF), conditional variational autoencoder (CVAE), and DNN, were trained by de Frahan and coworkers using filtered DNS data from the subvolume \mathcal{V}_3 of the low-swirl premixed flame, i.e., the algorithms were trained on \mathcal{D}_3^f , and the metrics were evaluated on \mathcal{D}_3^v (see Fig. 1). Figure 5 compares the marginal FDFs $P(Z)$ and $P(c)$ obtained using the three ML models, β -function model and DNS result for \mathcal{V}_3 for three different values (low, medium, and high) of the Jensen-Shannon divergence (JSD), which measures the similarity of two probability distributions, $Q_1 = Q^{\text{DNS}}(n)$ and $Q_2 = Q^{\text{model}}(n)$. The JSD is given by

$$J(Q_1 || Q_2) = \frac{1}{2} \sum_{n=1}^N \left\{ Q_1(n) \ln \left[\frac{Q_1(n)}{Q_2(n)} \right] + Q_2(n) \ln \left[\frac{Q_2(n)}{Q_1(n)} \right] \right\} \quad (19)$$

The JSD divergence is symmetric, i.e., $J(Q_1 || Q_2) = J(Q_2 || Q_1)$, and mathematically bounded between 0 and $\ln(2)$, with 0 indicating $Q_1 = Q_2$. The JSD for the three samples shown in Fig. 5 were computed by considering the FDFs extracted from the DNS of the premixed flame and those obtained by the $\beta - \beta$ analytical model. It can be seen from Fig. 5 that the $\beta - \beta$ analytical model is unable to capture more complex FDF shapes, such as bimodal distributions, as also confirmed by high JSD values. Thus, the need for more accurate models is motivated. Accurate predictions can be expected for $J(P || P_m) < 0.3$, whereas predictions with $J(P || P_m) > 0.6$ exhibit incorrect median values and overall shapes.

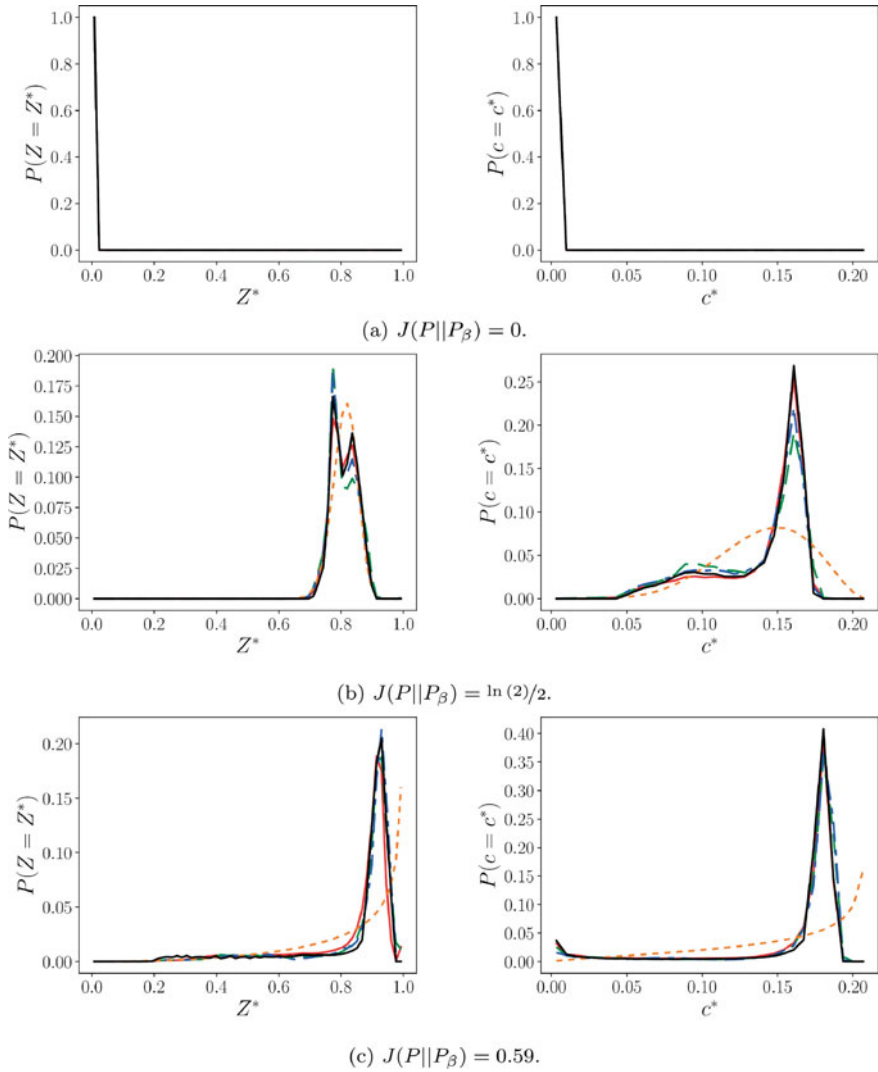


Fig. 5 Marginal FDF for low, mid-range, and high Jensen-Shannon divergence values for the $\beta - \beta$ PDF model. Red solid line is for RF model, green dashed line is for DNN model, blue dash-dotted line is for CVAE model, orange short dashed line is for $\beta - \beta$ model and black solid line is the DNS result (de Frahan et al. 2019)

The abilities of the three ML models to infer the subgrid FDF in regions other than \mathcal{D}_3^t was also assessed because DNS results showed that the FDF in downstream locations were significantly different from those for \mathcal{V}_3 . So, the ML models were trained using (1) \mathcal{D}_3^t data (volume centered at $z = 0.0775$ m), (2) data from \mathcal{D}_5^t (volume centered at $z = 0.1025$ m) and (3) data collected from the odd-numbered volumes $\mathcal{D}^t = \cup_{i=1,3,5,7,9} \mathcal{D}_i^t$. The training data in the last case were representative of the entire computational domain. It was found that the models trained using data from a single volume were unable to infer the FDF in other volumes which was indicated by the high 90th percentile (J_{90}) of all the Jensen–Shannon divergences errors. The ML models trained using the odd-numbered volumes (test 3 above) gave $J_{90} < 0.2$ for the entire physical domain although only 4% of the DNS data from the entire computational domain was used for the training. Among the three ML modes, DNN yielded the lowest errors. The analytical $\beta - \beta$ model had J_{90} values which were almost twice of that for the ML models. The sample marginal FDFs of mixture fraction and progress variable for 3 different values of Jensen–Shannon divergences computed for the DNN model are shown in Fig. 6 and it is clear that the bimodal distributions are also captured quite well by the ML models.

Another generalisation test was conducted by using validation data generated from a different time snapshot of the DNS ($t = 0.059$ s). For this case, the DNN model trained on $\mathcal{D}^t = \cup_{i=1,3,5,7,9} \mathcal{D}_i^t$ provided reasonable J_{90} values, although slightly higher than those obtained for the validation data from the same time snapshot of the training data. The $\beta - \beta$ model provided similar errors in both cases but three times higher than those of the DNN model. These generalisation tests demonstrated that the learned models are able to generalize temporally, as well as spatially. The results reported in this subsection suggest that it is important to use the training data covering the expected range of physical processes for which the ML is to be applied.

5.1.2 MILD Combustion

For the MILD combustion cases, the FDFs provided by DNN, $\beta - \beta$ and *copula* models are presented and compared to the DNS FDFs in Figs. 7, 8 and 9 for cases AZ1, AZ2 and BZ1 respectively. The DNN model significantly outperforms both analytical models and its prediction agrees very well with the DNS data for the different cases. As a general observation, the DNN captures the non-regular shapes of the marginal FDF of the progress variable quite well where the analytical models given by the β function and *copula* give *Gaussian*-like distributions. This difference has important implications for the reaction rate modelling as one shall see later in Sect. 5.2. For the mixture fraction, however, all models give good results but only the DNN is able to capture the asymmetry of the FDF which can be seen clearly in Fig. 9b and 9d for case BZ1. These results indicate promising capabilities of the DNN to predict the complex subgrid scalar statistics in MILD combustion.

It was noted by Chen et al. (2021) that the FDFs extracted directly using the instantaneous snapshots of DNS are random variables containing subgrid statistical information, as also pointed out in Pitsch (2006) and Pope (1985). The instantaneous

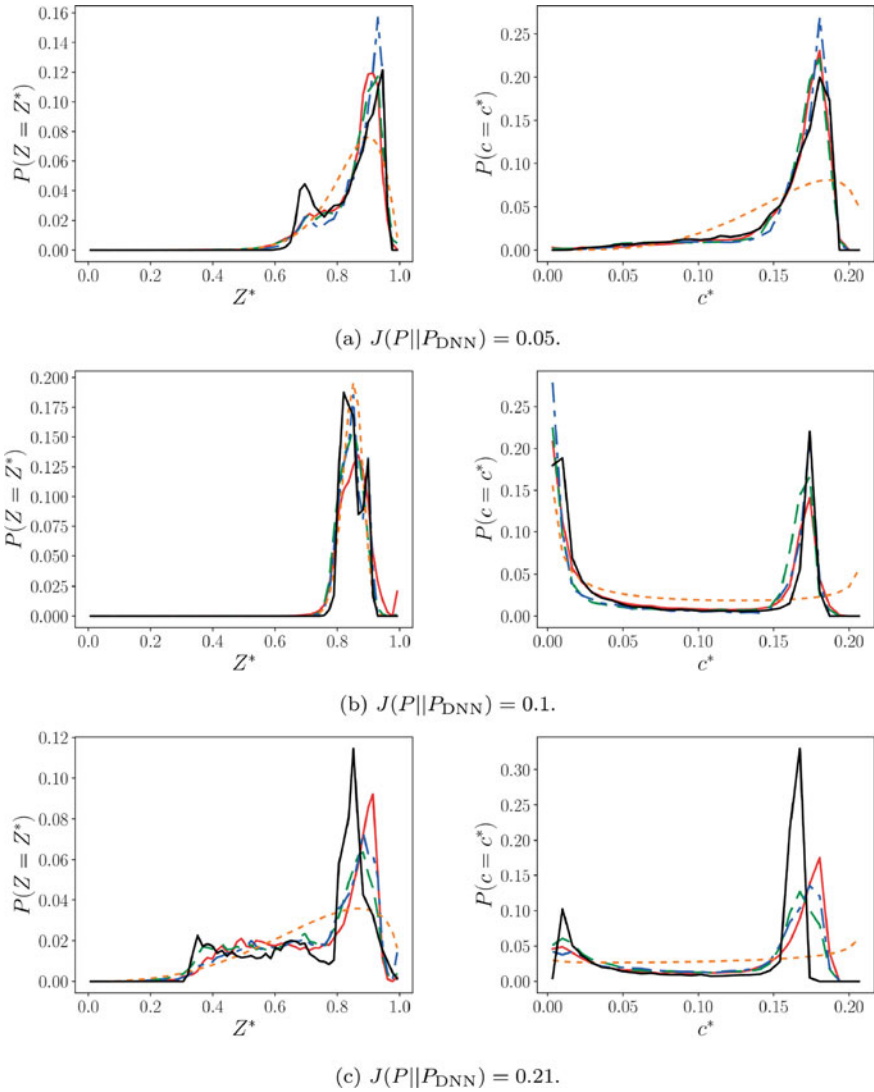


Fig. 6 Marginal FDF for median and high Jensen-Shannon divergence values for models trained on $\mathcal{D}^l = \cup_{i=1,3,5,7,9} D_i^l$. Red solid line is for RF, green dashed line is for DNN, blue dash-dotted line is for CVAE, orange short dashed line is for $\beta - \beta$ model, and black solid line is for DNS (de Frahan et al. 2019)

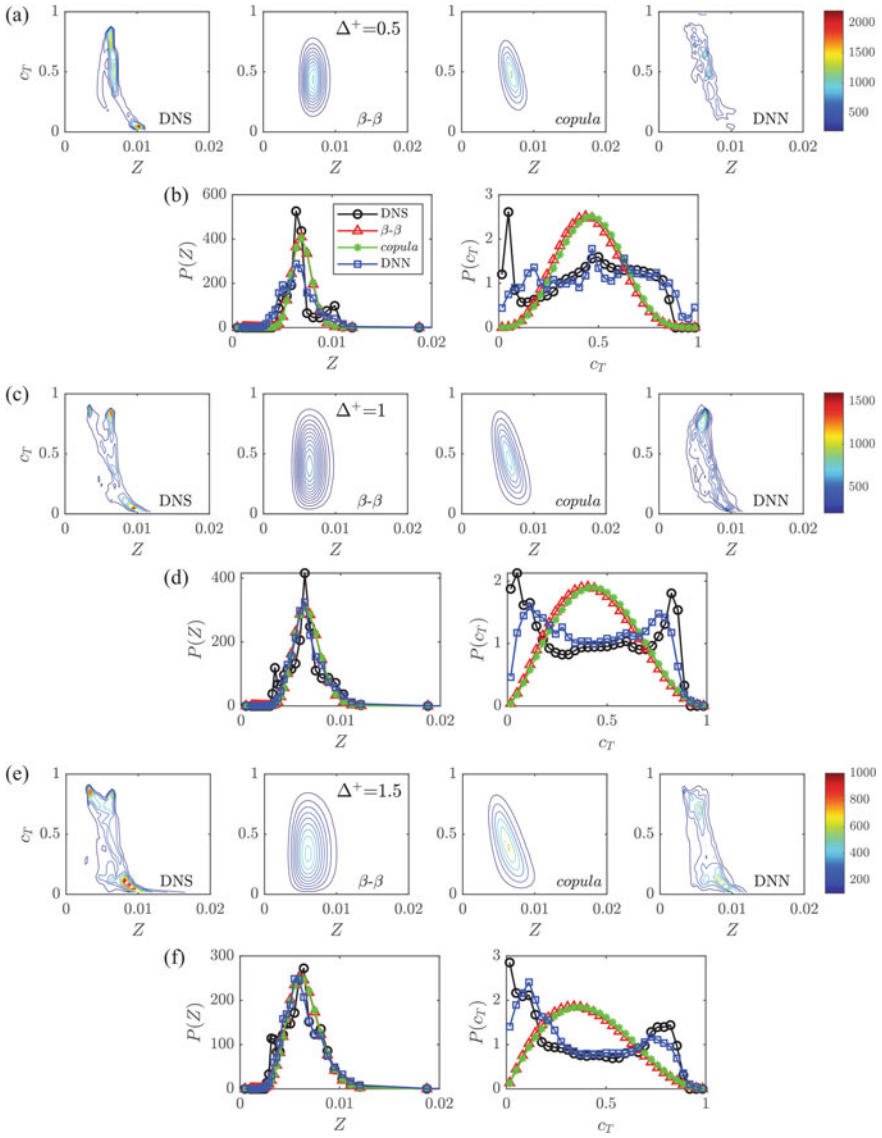


Fig. 7 Case AZ1: comparison of joint and marginal FDFs from DNS and models for filter sizes of $\Delta^+ = 0.5$ in (a) and (b), $\Delta^+ = 1$ in (c) and (d), and $\Delta^+ = 1.5$ in (e) and (f) (Chen et al. 2021)

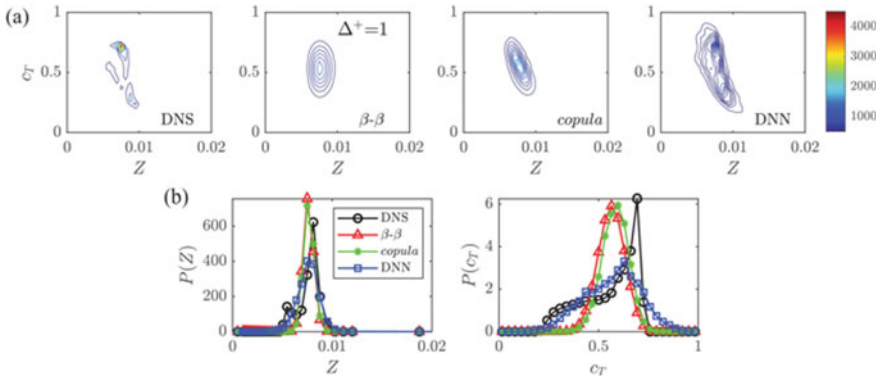


Fig. 8 Case AZ2: comparison of joint and marginal FDFs from DNS and models for filter sizes of $\Delta^+ = 0.5$ (Chen et al. 2021)

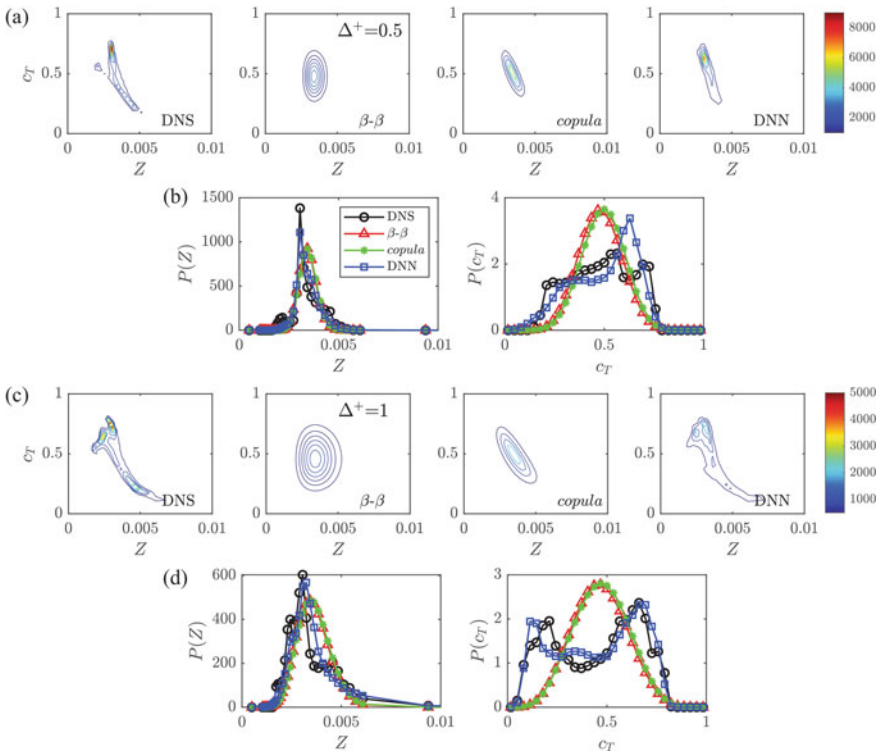


Fig. 9 Case BZ1: comparison of joint and marginal FDFs from DNS and models for filter sizes of $\Delta^+ = 0.5$ in (a) and (b), and $\Delta^+ = 1.0$ in (c) and (d) (Chen et al. 2021)

FDFs present certain levels of randomness due to the unsteady nature of single realisations. This randomness is removed to a good extent if the training data for ML are selected over many DNS realisations at a statistically stationary state. Therefore, following several experimental studies (Wang et al. 2007; Tong 2001; Cai et al. 2009), the instantaneous FDFs obtained from the DNS were conditioned on the resolved scalars, \tilde{Z} and \tilde{c}_T , and then ensemble-averaged. A quantitative comparison of the conditionally averaged FDFs was then performed. Two variables, \tilde{Z} and \tilde{c}_T , were considered as the number of available DNS samples was not sufficient to perform a statistically meaningful averaging on the four statistical moments used as ANN inputs. The resolved mixture fraction and progress variable were chosen so that the selected samples were located in the reaction zone ($\tilde{c}_T \approx 0.5$). Figures 10 and 11 show the conditional FDFs, $\langle \tilde{P}(Z, c_T) | \tilde{Z}, \tilde{c}_T \rangle$, for cases AZ1 and BZ1 respectively and the values of the conditioning variables are given in the figure captions. The DNN accurately reproduces the conditional joint and both marginal FDFs. It also captures the significant changes in the FDF shape with the varying filter size, especially for the progress variable. For case AZ1, both the β and *copula* models overpredict the peak when $\Delta^+ \leq 1$ for both Z and c_T distributions. However, for $\Delta^+ = 1.5$, the overall prediction is good for $\tilde{P}(Z)$ and the peak of $\tilde{P}(c_T)$ is also close to the DNS value although the shape is not captured. Similar results were reported for cases AZ2 also. For case BZ1, the mixture fraction distribution is predicted fairly well by all models for different Δ^+ values. However, both analytical models fail to predict the *bimodal-plateau* shape of $\tilde{P}(c_T)$, which is typical of MILD combustion but seen seldom in conventional flames.

The JSD values were also calculated using Eq. (19), for the DNN and the two analytical models which confirmed the observations made using Figs. 7, 8, 9, 10 and 11. The JSD values provided by the DNN were much lower than those for the β and *copula* models. Improved predictions and lower JSD values were observed for all the models by increasing the filter size and this improvement was particularly significant for the DNN having $J_{90} < 0.05$. The DNN model performed equally well for Z and c_T .

To check for generalisation capability, the DNN was further validated using data which were not included in the learning/training step. The training and validation datasets included snapshots taken from $t = \tau_f$ to $1.2\tau_f$, where τ_f is the flow-through time, but the test data were taken using snapshots taken between $1.4\tau_f$ and $1.5\tau_f$. Substantial variations in the MILD combustion behaviour were observed among these snapshots (see Doan et al. 2018 for details). Hence, a robustly trained DNN is attractive if it can accurately infer a quantity of interest (here, FDF) for scenarios that have not been explicitly *seen* during the training process. The PDFs of the JSD values for the self-predictions (i.e., predictions performed on the training datasets) and unknown-predictions of the FDF are shown in Fig. 12. A filter size of $\Delta^+ = 1$ was used for all cases. As indicated in Fig. 12, the DNN provides a similar level of accuracy when *unseen* test data points are fed to the model. More than 80% of the JSD values are smaller than 0.05. The advantage of using DNN as FDF model is still unaffected since the majority of JSD values were larger than 0.1 for the β and *copula* FDF models. A slightly worse performance was achieved by the DNN when

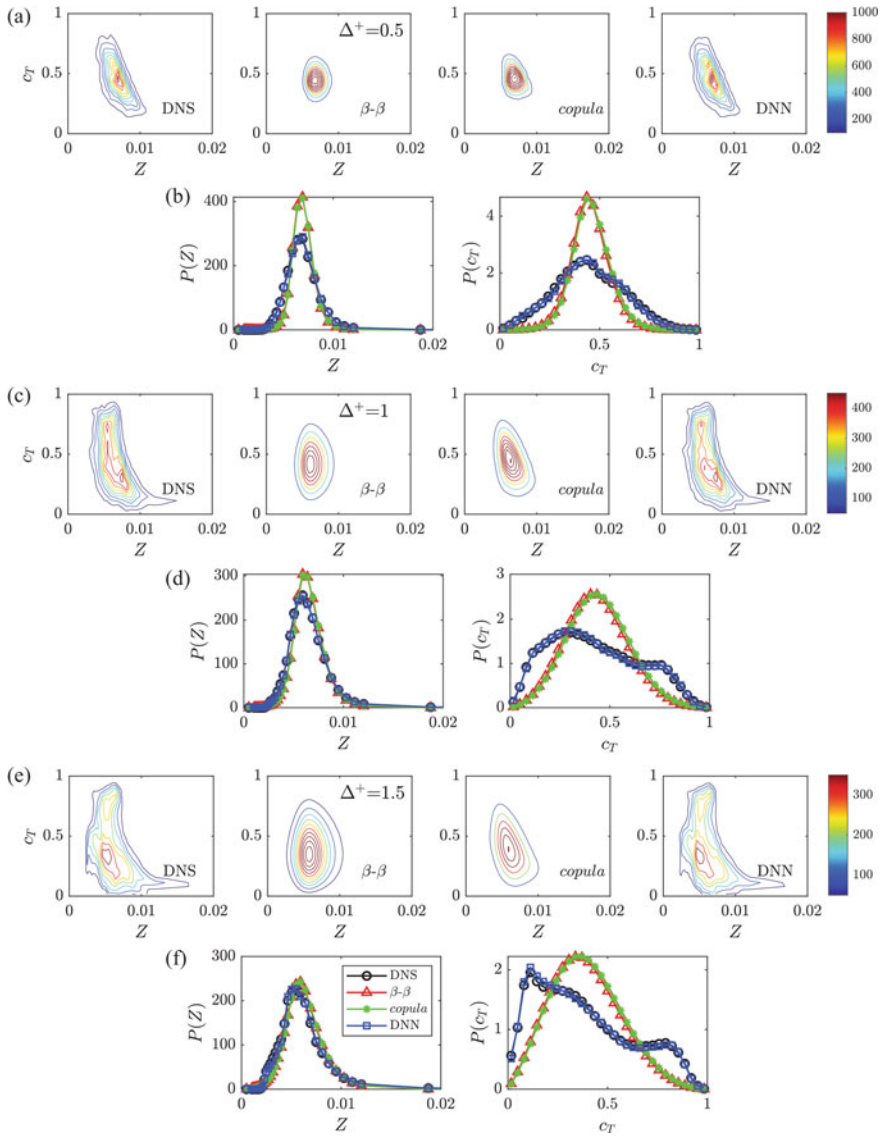


Fig. 10 Case AZ1: comparison of joint and marginal FDFs from DNS and models for **a** and **b** $\Delta^+ = 0.5$, $\tilde{Z} = 0.007$, $\tilde{c}_T = 0.45$; **c** and **d** $\Delta^+ = 1$, $\tilde{Z} = 0.0066$, $\tilde{c}_T = 0.43$; and **e** and **f** $\Delta^+ = 1.5$, $\tilde{Z} = 0.0064$, $\tilde{c}_T = 0.39$ (Chen et al. 2021)

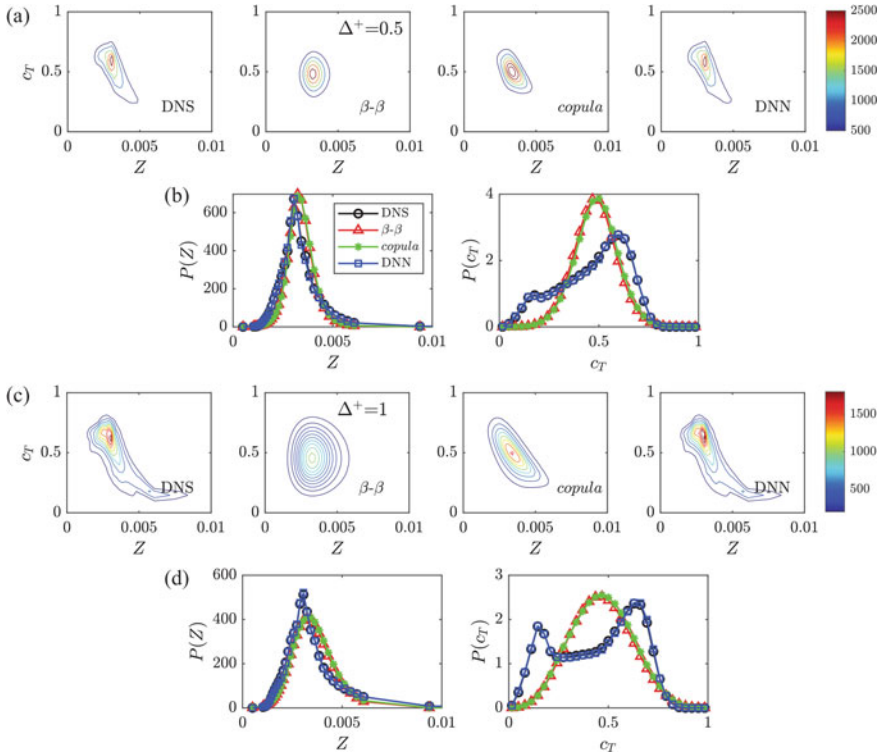


Fig. 11 Case BZ1: comparison of joint and marginal FDFs from DNS and models for **a** and **b** $\Delta^+ = 0.5$, $\tilde{Z} = 0.00034$, $\tilde{c}_T = 0.48$; and **c** and **d** $\Delta^+ = 1$, $\tilde{Z} = 0.0036$, $\tilde{c}_T = 0.46$ (Chen et al. 2021)

the training data came from cases AZ1 and BZ1, and the validation was done on case AZ2. The JSD results obtained from this new test with the self-predictions for $\Delta^+ = 0.5$ indicated that the overall performance was still good although the JSD distribution shifted towards higher JSD values. Further improvement on predictions is expected to be achieved if more datasets with different scenarios are included in the training.

5.1.3 Spray Flame

Yao et al. (2020) visually compared the FDF predicted by ANN and β -function model with the DNS values for one of the validation cases (CX1). Moreover, the data samples of this case were divided into three different groups characterized by filter size Δ_{LES} , to compute the sensitivity of the trained ANN model to LES grid sizes. The LES cells were selected randomly for a given $\tilde{\xi}$ ranging from fuel-lean to fuel-rich conditions. The stoichiometric mixture fraction value is $\tilde{\xi}_{st} = 0.068$.

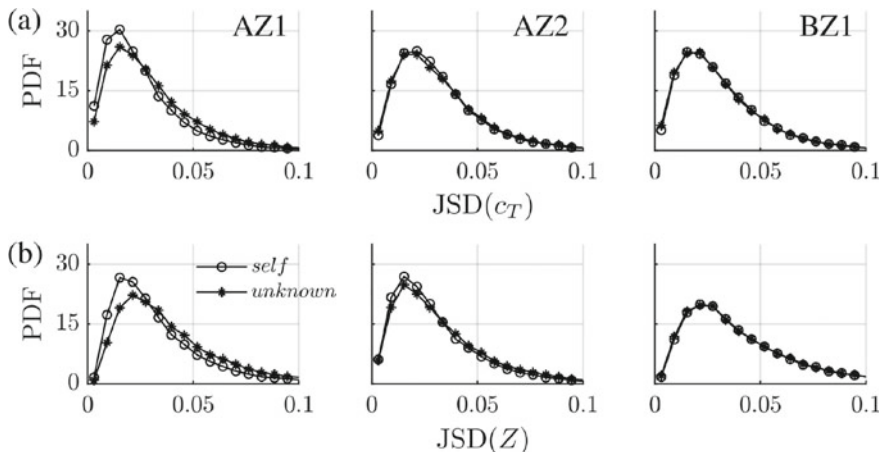


Fig. 12 Comparison of Jansen-Shannon divergence for DNN self- and unknown-predictions of FDF of **a** progress variable and **b** mixture fraction. The filter size for all cases is $\Delta^+ = 1.0$ (Chen et al. 2021)

Figure 13 compares FDF computed using ANN and β -function with DNS results for two filtered mixture fraction values and three Δ_{LES} . There is no marked differences in the ANN prediction for different Δ_{LES} . The ANN predictions of $\tilde{P}(\eta)$ are in excellent agreement with the DNS results, including the peak value and its location. The FDF is skewed towards the lean side ($\eta < \xi_{st}$) for $\tilde{\xi} = 0.05$ whereas it is stretched towards the rich side for $\tilde{\xi} = 0.10$, and even a bimodal behaviour appears at larger filter sizes. The β -function does not seem to represent the FDFs well and numerical issues can arise when the mean is close to zero or unity with small SGS variance (Kronenburg et al. 2000).

5.2 Reaction Rate Predictions

The filtered reaction rate inferred by the ML models were also assessed against DNS results by de Frahan et al. (2019) for their premixed flame and by Chen et al. (2021) for the MILD combustion cases. The ML models used by de Frahan et al. inferred the unconditional filtered reaction rates $\bar{\omega}$, which are computed according to Eq. 6, and are shown in Fig. 14. Significant over predictions were observed for the $\beta - \beta$ model. The comparisons of the conditional reaction rates are also shown in Fig. 14.

The reaction rate in the transport equation for the filtered temperature-based progress variable, $\bar{\omega}_{c_T}$, can be computed using

$$\bar{\omega}_{c_T}(\mathbf{x}, t) = \int_0^1 \int_0^1 \langle \dot{\omega}_{c_T} \rangle \tilde{P}(Z, c_T; \mathbf{x}, t) dZ dc_T, \quad (20)$$

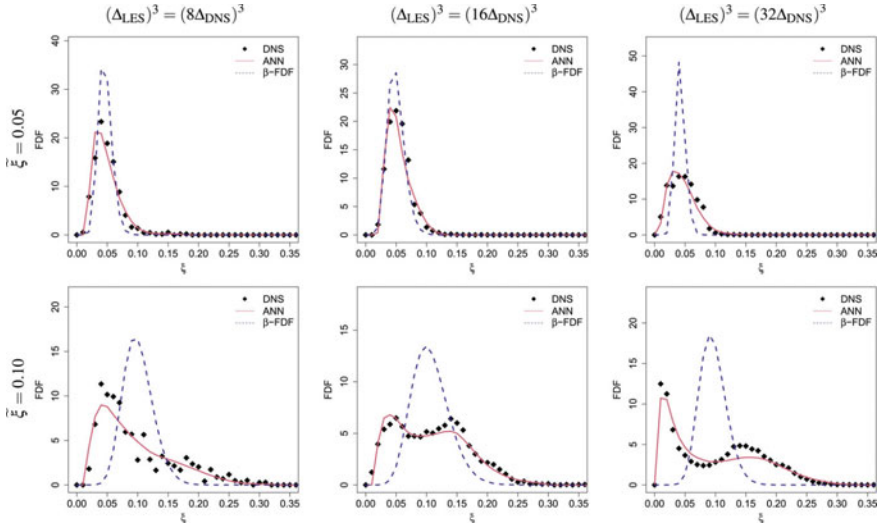


Fig. 13 Validation of ANN predictions of $\tilde{P}(\eta)$ with DNS results for different LES grid sizes. The results are shown for $\tilde{\xi} = 0.05$ (top) and $\tilde{\xi} = 0.1$ (bottom) (Yao et al. 2020)

where the joint FDF $\tilde{P}(Z, c_T)$ is obtained through the ANN in the MILD combustion cases investigated by Chen et al. (2021). The symbol $\langle \dot{\omega}_{c_T}(\mathbf{x}, t) \rangle = \langle \dot{\omega}_{c_T}(\mathbf{x}, t) / \rho(\mathbf{x}, t) | Z, c_T \rangle$ is defined as the doubly conditional mean reaction rate obtained from the DNS data. The instantaneous reaction rate of c_T is defined as $\dot{\omega}_{c_T} = \dot{q} / [c_p(T_b - T_u)]$, with \dot{q} and c_p being the volumetric heat release rate and specific heat capacity of the mixture respectively. The conditional averages are computed using samples collected over the entire computational domain, see Sect. 3.2, and all the snapshots available (≈ 60) to achieve good statistical convergence. The authors verified that the doubly conditional mean rates have negligible variations in time and space, supporting the assumption of many turbulent combustion models (*viz.*, flamelets, see Bradley et al. 1990; Fiorina et al. 2003; Pierce and Moin 2004; van Oijen et al. 2016; and conditional moment-based methods, see Klimenko and Bilger 1999; Steiner and Bushe 2001) that the conditional means have small temporal and spatial variations if appropriate conditioning variables are used. The target filtered reaction rate $\overline{\dot{\omega}_{c_T}}^{m-DNS}$ was obtained by computing both the conditional mean reaction rate and the FDF in Eq. 20 directly from the DNS data. The scatter plots of $\overline{\dot{\omega}_{c_T}}^{m-DNS}$ and the reaction rates computed using FDFs obtained through β , *copula* and DNN models are presented in Fig. 15 for one of the DNS cases (AZ1) investigated in Chen et al. (2021). The qualitative behaviours and the trends were found to be similar for the other two cases. Although all models give reasonable predictions, the DNN outperforms the analytical models for all filter sizes. Moreover, the DNN predictions generally exhibit good symmetry about the diagonal, indicating a bias towards neither under- nor over-prediction, while the scatters for both the β

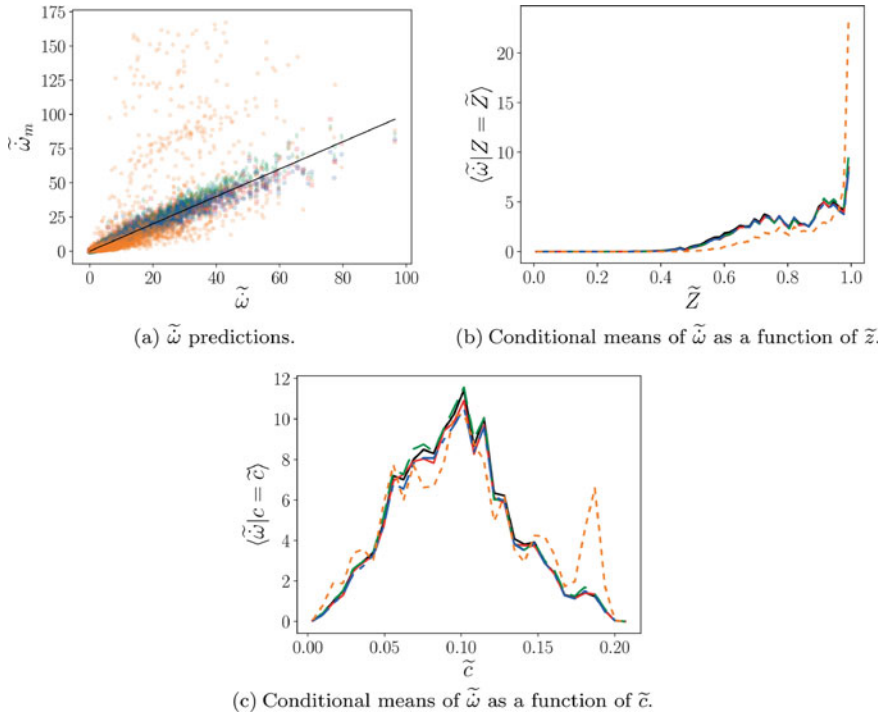


Fig. 14 Reaction rate $\tilde{\omega}$ inferred by the ML models trained on $\mathcal{D}^f = \cup_{i=1,3,5,7,9} D_i^f$. Red squares and solid line are for RF model, green diamonds and dashed line are for DNN, blue circles and dash-dotted line are for CVAE, orange pentagons and short dashed line are for $\beta - \beta$ model, and black solid line is for DNS result (de Frahan et al. 2019)

and *copula* models are asymmetric. As Δ^+ increases, the DNN prediction improves considerably whereas the performance of the analytical models does not follow this trend with the filter size. For both the β and *copula* models, a trend in the off-diagonal samples moving from under-predictions at small Δ^+ to over-predictions at larger Δ^+ can be seen.

6 Conclusions and Prospects

The application of ML algorithms to infer subgrid-scale filtered density functions (FDFs) in three test cases, i.e., swirling premixed flame, MILD and spray combustion, have been discussed in this chapter. Particularly, the promising results provided by deep neural networks (DNNs) for accurately inferring the FDFs have been shown. DNNs are generally able to capture the complex FDF behaviours and their variations with great accuracy across various combustion scenarios, turbulent and thermochem-

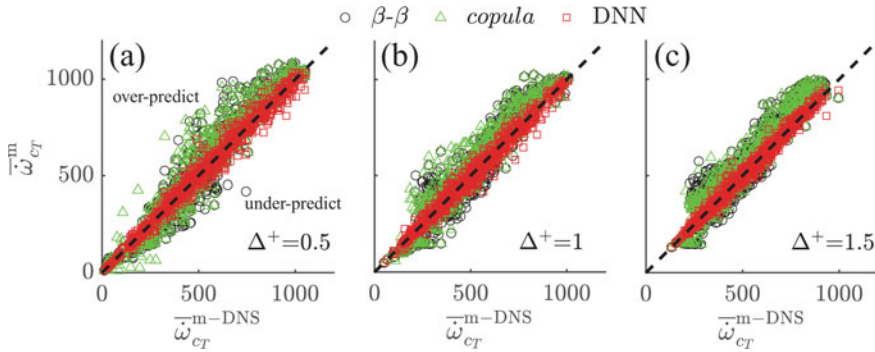


Fig. 15 Scatter plot of $\bar{\omega}_{c_T}^{m-DNS}$ and $\bar{\omega}_{c_T}$ (in $\text{kg}/\text{m}^3/\text{s}$) modelled using different FDF models (denoted using different markers) for case AZ1. The results for different filter sizes are also shown (Chen et al. 2021)

ical conditions, and LES filter sizes. This can be achieved by manipulating the input data (extracted from DNS of these three cases), changing the network architecture, and tuning the network hyperparameters (e.g., learning rate, batch size). It has been shown that if the DNN training dataset is heterogeneous, i.e., it contains different possible outcomes of the quantities of interest, the DNN can handle *unknown* inputs quite well, suggesting a good model robustness. Thus, the DNN can be applied as a *black-box* model to other cases. By contrast, analytical models such as the β -function and *copula* models in most cases show their limitations quite clearly.

Although the above observations demonstrate the potential of DNN-based FDF modelling in combustion, several challenges remain and require further investigations. Searching for an optimal combination of the DNN hyperparameters can be highly time-consuming and computationally expensive. For example, an exhaustive grid search, looping through all combinations of layers and neurons to find an optimum, is not an easy task and may require cloud computing services (Yao et al. 2020). Moreover, due to the black-box nature of ML models, it is often hard to debug them to a satisfying level or improve them substantially after such a level is reached. This shifts the attention to the preprocessing of training data, which can be a daunting and time-consuming task, as mentioned in Chen et al. (2021). The lack of physical constraints in the training of ML models is yet another issue, and research is ongoing to develop physics-informed ML models that can respect physical laws and increase the interpretability and generalisation capability of ML models.

If DNNs are to replace combustion models, the overhead of retrieving predictions can also be of concern and counterbalance the observed savings in storage requirement. The overhead associated with the use of DNNs is highly machine-dependent and also network size-dependent. A posteriori LES studies need to quantify the computational times required by the DNN inference of FDFs and mean reaction rates. High inference times could hinder the development of in-situ capabilities, where the ML model is trained during the simulation, which can mitigate the risk of extrapola-

tion. The latter can be reduced by also combining ML training and applications with uncertainty quantification or sensitivity analysis approaches that can effectively verify the performance of the ML model, provide a level of confidence in its predictions, guarantee that it does not violate physics laws and promote its more comprehensive application.

Machine Learning has induced notable advancements in combustion science. It has been effectively used for finding hidden patterns under large amounts of data, exploring and visualising high-dimensional input spaces, deriving complex mapping from inputs and outputs, and reducing computational cost and memory occupation (Zhou et al. 2022). However, many challenges and hence research opportunities are left to be addressed, and the development of physics-based ML approaches is just the starting point of a scientific paradigm shift that will bring new insights in combustion science with the help of big data. The combination of ML and combustion will provide solutions to daunting problems and enhance the understanding and deployment of novel combustion processes and technologies, which will shape a cleaner and sustainable future energy arena.

Acknowledgements Z. Li and N. Swaminathan acknowledge the support from EPSRC through the grant EP/S025650/1. Iavarone acknowledges the support of FRS-FNRS Fellowship.

References

- An J, He G, Luo K, Qin F, Liu B (2020) Artificial neural network based chemical mechanisms for computationally efficient modeling of hydrogen/carbon monoxide/kerosene combustion. *Int J Hydrogen Energy* 45(53):29594–29605
- Ali Sen B, Menon S (2010) Linear eddy mixing based tabulation and artificial neural networks for large eddy simulations of turbulent flames. *Combust Flame* 157(1):62–74
- Bilger RW (1976) Structure of diffusion flames. *Combust Sci Technol* 13:155–170
- Blasco JA, Fueyo N, Dopazo C, Ballester J (1998) Modelling the temporal evolution of a reduced combustion chemical system with an artificial neural network. *Combust Flame* 113(1–2):38–52
- Blasco JA, Fueyo N, Larroya JC, Dopazo C, Chen J-Y (1999) A single-step time-integrator of a methane-air chemical system using artificial neural networks. *Comp Chem Eng* 23(9):1127–1133
- Blasco J, Fueyo N, Dopazo C, Chen J-Y (2000) A self-organizing-map approach to chemistry representation in combustion applications. *Combust Theory Model* 4(1):61–76
- Bradley D, Gaskell PH, Lau AKC (1990) A mixedness-reactedness flamelet model for turbulent diffusion flames. *Proc Combust Inst* 23(1):685–692
- Bradley D, Gaskell PH, Gu XJ (1998) The mathematical modeling of liftoff and blowoff of turbulent non-premixed methane jet flames at high strain rates. *Proc Combust Inst* 27(1):1199–1206
- Cai J, Wang D, Tong C, Barlow RS, Karpets AN (2009) Investigation of subgrid-scale mixing of mixture fraction and temperature in turbulent partially premixed flames. *Proc Combust Inst* 32(1):1517–1525
- Chatzopoulos AK, Rigopoulos S (2013) 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(1):1465–1473
- Chen J-Y, Blasco JA, Fueyo N, Dopazo C (2000) An economical strategy for storage of chemical kinetics: fitting in situ adaptive tabulation with artificial neural networks. *Proc Combust Inst* 28(1):115–121

- Chen ZX, Ruan S, Swaminathan N (2015) Simulation of turbulent lifted methane jet flames: effects of air-dilution and transient flame propagation. *Combust Flame* 162:703–716
- Chen ZX, Ruan S, Swaminathan N (2017) Large eddy simulation of flame edge evolution in a spark-ignited methane-air jet. *Proc Combust Inst* 36:1645–1652
- Chen ZX, Doan NAK, Ruan S, Langella I, Swaminathan N (2018) A priori investigation of subgrid correlation of mixture fraction and progress variable in partially premixed flames. *Combust Theory Model* 22:862–882
- Chen ZX, Iavarone S, Ghiasi G, Kannan V, D'Alessio G, Parente A, Swaminathan N (2021) Application of machine learning for filtered density function closure in MILD combustion. *Combust Flame* 225:160–179
- Chi C, Janiga G, Thévenin D (2021) On-the-fly artificial neural network for chemical kinetics in direct numerical simulations of premixed combustion. *Combust Flame* 226:467–477
- Christo FC, Masri AR, Nebot EM, Pope SB (1996a) An integrated pdf/neural network approach for simulating turbulent reacting systems. *Symp (International) Combust* 26(1):43–48
- Christo FC, Masri AR, Nebot EM, Turanyi T (1995) Utilising artificial neural network and reparamodelling in turbulent combustion. In: *Proceedings of ICNN'95 - international conference on neural networks*, vol 2. IEEE, pp 911–916
- Christo FC, Masri AR, Nebot EM (1996) Artificial neural network implementation of chemistry with pdf simulation of h₂/co₂ flames. *Combust Flame* 106(4):406–427
- Cook AW, Riley JJ (1994) A subgrid model for equilibrium chemistry in turbulent flows. *Phys Fluids* 6:2868–2870
- Darbyshire OR, Swaminathan N (2012) A presumed joint pdf model for turbulent combustion with varying equivalence ratio. *Combust Sci Technol* 184(12):2036–2067
- Day M, Tachibana S, Bell J, Lijewski M, Beckner V, Cheng RK (2012) A combined computational and experimental characterization of lean premixed turbulent low swirl laboratory flames: I. methane flames. *Combust Flame* 159(1):275–290
- de Frahan MTH, Yellapantula S, King R, Day MS, Grout RW (2019) Deep learning for presumed probability density function models. *Combust Flame* 208:436–450
- Ding T, Readshaw T, Rigopoulos S, Jones WP (2021) Machine learning tabulation of thermochemistry in turbulent combustion: an approach based on hybrid flamelet/random data and multiple multilayer perceptrons. *Combust Flame* 231:111493
- Doan NAK, Swaminathan N, Minamoto Y (2018) DNS of MILD combustion with mixture fraction variations. *Combust Flame* 189:173–189
- Emami MD, Fard AE (2012) Laminar flamelet modeling of a turbulent ch₄/h₂/n₂ jet diffusion flame using artificial neural networks. *App Math Model* 36(5):2082–2093
- Fiorina B, Baron R, Gicquel O, Thevenin D, Carpentier S, Darabiha N (2003) Modelling non-adiabatic partially premixed flames using flame-prolongation of ILDM. *Combust Theory Model* 7:449–470
- Flemming F, Sadiki A, Janicka J (2005) LES using artificial neural networks for chemistry representation. *Prog Comput Fluid Dyn* 5(7):375–385
- Fox RO (2003) *Computational models for turbulent reacting flows*. Cambridge University Press, Cambridge, UK
- Franke LLC, Chatzopoulos AK, Rigopoulos S (2017) Tabulation of combustion chemistry via artificial neural networks (anns): methodology and application to LES-PDF simulation of Sydney flame I. *Combust Flame* 185:245–260
- Gicquel O, Darabiha N, Thevenin D (2000) Laminar premixed hydrogen/air counterflow flame simulations using flame prolongation of ILDM with differential diffusion. *Proc Combust Inst* 28(2):1901–1908
- Giusti A, Sitte MP, Borghesi G, Mastorakos E (2018) Numerical investigation of kerosene single droplet ignition at high-altitude relight conditions. *Fuel* 225:663–670
- Goodfellow I, Bengio Y, Courville A (2016) *Deep learning*. MIT Press
- Grout RW, Swaminathan N, Cant RS (2009) Effects of compositional fluctuations on premixed flames. *Combust Theory Model* 13(5):823–852

- Ihme M, Pitsch H (2008a) Prediction of extinction and reignition in nonpremixed turbulent flames using a flamelet/progress variable model: 1. a priori study and presumed pdf closure. *Combust Flame* 155(1):70–89
- Ihme M, Pitsch H (2008b) Prediction of extinction and reignition in nonpremixed turbulent flames using a flamelet/progress variable model: 2. application in LES of Sandia flames D and E. *Combust Flame* 155(1-2):90–107
- Ihme M, Marsden AL, Pitsch H (2006) On the optimization of artificial neural networks for application to the approximation of chemical systems. Center for Turbulence Research Annual Research Briefs, pp 105–118
- Ihme M, Marsden AL, Pitsch H (2008) Generation of optimal artificial neural networks using a pattern search algorithm: application to approximation of chemical systems. *Neural Comput* 20(2):573–601
- Ihme M, Schmitt C, Pitsch H (2009) Optimal artificial neural networks and tabulation methods for chemistry representation in les of a bluff-body swirl-stabilized flame. *Proc Combust Inst* 32(1):1527–1535
- Ihme M, Chung WT, Mishra AA (2022) Combustion machine learning: principles, progress and prospects. *Prog Energy Combust Sci* 91:101010. <https://doi.org/10.1016/j.pecs.2022.101010>
- Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. [arXiv:1502.03167](https://arxiv.org/abs/1502.03167)
- Jin B, Grout R, Bushe WK (2008) Conditional source-term estimation as a method for chemical closure in premixed turbulent reacting flow. *Flow Turbul Combust* 81(4):563–582
- Jones WP, Kakhi M (1998) Pdf modeling of finite-rate chemistry effects in turbulent nonpremixed jet flames. *Combust Flame* 115(1–2):210–229
- Jordan MI, Mitchell TM (2015) Machine learning: trends, perspectives, and prospects. *Science* 349(6245):255–260
- Kempf A, Flemming F, Janicka J (2005) Investigation of lengthscales, scalar dissipation, and flame orientation in a piloted diffusion flame by les. *Proc Combust Inst* 30(1):557–565
- Khan QS, Baek SW, Ghassemi H (2007) On the autoignition and combustion characteristics of kerosene droplets at elevated pressure and temperature. *Combust Sci Technol* 179(12):2437–2451
- Kingma DP, Ba J (2014) Adam: a method for stochastic optimization. [arXiv: 1412.6980](https://arxiv.org/abs/1412.6980)
- Klimenko AY, Bilger RW (1999) Conditional moment closure for turbulent combustion. *Prog Energy Combust Sci* 25(6):595–687
- Kronenburg A, Bilger RW (2000) Kent JH (2000) Computation of conditional average scalar dissipation in turbulent jet diffusion flames. *Flow Turbul Combust* 64(3):145–159
- Linse D, Kleemann A, Hasse C (2014) Probability density function approach coupled with detailed chemical kinetics for the prediction of knock in turbocharged direct injection spark ignition engines. *Combust Flame* 161(4):997–1014
- Minamoto Y, Swaminathan N, Cant RS, Leung T (2014) Reaction zones and their structure in MILD combustion. *Combust Sci Technol* 186(8):1075–1096
- Minamoto Y, Swaminathan N (2015) Subgrid scale modelling for MILD combustion. *Proc Combust Inst* 35(3):3529–3536
- Minamoto Y, Swaminathan N (2014) Scalar gradient behaviour in MILD combustion. *Combust Flame* 161(4):1063–1075
- Navarro-Martinez S, Kronenburg A, Di Mare F (2005) Conditional moment closure for large eddy simulations. *Flow Turbul Combust* 75(1):245–274
- O’Brien EE, Jiang TL (1991) The conditional dissipation rate of an initially binary scalar in homogeneous turbulence. *Phys Fluids A* 3(12):3121–3123
- Owoyele O, Kundu P, Ameen MM, Echehki T, Som S (2020) Application of deep artificial neural networks to multi-dimensional flamelet libraries and spray flames. *Int J Engine Res* 21(1):151–168
- Pierce CD, Moin P (2004) Progress-variable approach for large-eddy simulation of non-premixed turbulent combustion. *J Fluid Mech* 504:73–97

- Pitsch H (2006) Large-Eddy simulation of turbulent combustion. *Annu Rev Fluid Mech* 38:453–482
- Plackett RL (1965) A class of bivariate distributions. *J Amer Stat Assoc* 310:516–522
- Pope SB (2000) *Turbulent flows*. Cambridge University Press, Cambridge
- Pope SB (1985) Pdf methods for turbulent reactive flows. *Prog Energy Combust Sci* 11:119–192
- Pope SB (1990) Computations of turbulent combustion: progress and challenges. *Proc Combust Inst* 23:591–612
- Pope SB (2013) Small scales, many species and the manifold challenges of turbulent combustion. *Proc Combust Inst* 34:1–31
- Raman V, Pitsch H, Fox RO (2005) Hybrid large-eddy simulation/lagrangian filtered-density-function approach for simulating turbulent combustion. *Combust Flame* 143(1):56–78
- Ranade R, Li G, Li S, Echehki T (2021) An efficient machine-learning approach for pdf tabulation in turbulent combustion closure. *Combust Sci Technol* 193(7):1258–1277
- Readshaw T, Ding T, Rigopoulos S, Jones WP (2021) Modeling of turbulent flames with the large eddy simulation-probability density function (LES-PDF) approach, stochastic fields, and artificial neural networks. *Phys. Fluids* 33(3)
- Ruan S, Swaminathan N, Darbyshire OR (2014) Modelling of turbulent lifted jet flames using flamelets: a priori assessment and a posteriori validation. *Combust Theory Model* 18(2):295–329
- Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533–536
- Sen BA, Menon S (2009) Turbulent premixed flame modeling using artificial neural networks based chemical kinetics. *Proc Combust Inst* 32(1):1605–1611
- Sen BA, Hawkes ER, Menon S (2010) Large eddy simulation of extinction and reignition with artificial neural networks based chemical kinetics. *Combust Flame* 157(3):566–578
- Steiner H, Bushe WK (2001) Large eddy simulation of a turbulent reacting jet with conditional source-term estimation. *Phys Fluids* 13(3):754–769
- Stineman RW (1980) A consistently well-behaved method for interpolation. *Creative Comput* 6:54–57
- Swaminathan N, Bai X-S, Haugen NEL, Fureby C, Brethouwer G (eds) (2022) *Advanced turbulent combustion physics and applications*. Cambridge University Press, Cambridge, UK
- Tong C (2001) Measurements of conserved scalar filtered density function in a turbulent jet. *Phys Fluids* 13:2923
- van Oijen JA, Donini A, Bastiaans RJM, ten Thije Boonkkamp JHM, de Goey LPH (2016) State-of-the-art in premixed combustion modeling using flamelet generated manifolds. *Prog Energy Combust Sci* 57:30–74
- van Oijen JA, de Goey LPH (2002) Modelling of premixed counterflow flames using the flamelet-generated manifold method. *Combust Theory Model* 6:463–478
- Verdonck T, Hubert M, Rousseeuw P (2009) Robust PCA for skewed data and its outlier map. *Comput Stat Data Anal* 53:2264–2274
- Veynante D, Vervisch L (2002) Turbulent combustion modeling. *Prog Energy Combust Sci* 28:193–266
- Wan K, Barnaud C, Vervisch L, Domingo P (2020) 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* 220:119–129
- Wan K, Barnaud C, Vervisch L, Domingo P (2021) Machine learning for detailed chemistry reduction in dns of a syngas turbulent oxy-flame with side-wall effects. *Proc Combust Inst* 38(2):2825–2833
- Wang D, Tong C, Barlow RS, Karpets AN (2007) Experimental study of scalar filtered mass density function in turbulent partially premixed flames. *Proc Combust Inst* 31(1):1533–1541
- Wang F, Liu R, Li M, Yao J, Jin J (2018) Kerosene evaporation rate in high temperature air stationary and convective environment. *Fuel* 211:582–590
- Wang B, Kronenburg A, Stein OT (2019) A new perspective on modelling passive scalar conditional mixing statistics in turbulent spray flames. *Combust Flame* 208:376–387

- Wright YM, De Paola G, Boulouchos K, Mastorakos E (2005) Simulations of spray autoignition and flame establishment with two-dimensional cmc. *Combust Flame* 143(4):402–419
- Yao S, Wang B, Kronenburg A, Stein OT (2020) Modeling of sub-grid conditional mixing statistics in turbulent sprays using machine learning methods. *Phys Fluids* 32(11)
- Zhang Y, Xu S, Zhong S, Bai X-S, Wang H, Yao M (2020) Large eddy simulation of spray combustion using flamelet generated manifolds combined with artificial neural networks. *Energy AI* 2:100021
- Zheng Z, Lin X, Yang M, He Z, Bao E, Zhang H, Tian Z (2020) Progress in the application of machine learning in combustion studies. *Energy Environ* 9:1–14
- Zhou ZJ, Lü Y, Wang ZH, Xu YW, Zhou JH, Cen K (2013) Systematic method of applying ann for chemical kinetics reduction in turbulent premixed combustion modeling. *Chinese Sci Bul* 58(4):486–492
- Zhou L, Song Y, Ji W, Wei H (2022) Machine learning for combustion. *Energy AI* 7:100128

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

