



**HAL**  
open science

# Generalization Capability of Convolutional Neural Networks for Progress Variable Variance and Reaction Rate Subgrid-Scale Modeling

Victor Xing, Corentin Lapeyre, Thomas Jaravel, Thierry Poinso

► **To cite this version:**

Victor Xing, Corentin Lapeyre, Thomas Jaravel, Thierry Poinso. Generalization Capability of Convolutional Neural Networks for Progress Variable Variance and Reaction Rate Subgrid-Scale Modeling. *Energies*, 2021, 14 (16), pp.5096. 10.3390/en14165096 . hal-04538176

**HAL Id: hal-04538176**

**<https://ut3-toulouseinp.hal.science/hal-04538176v1>**

Submitted on 9 Apr 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article

# Generalization Capability of Convolutional Neural Networks for Progress Variable Variance and Reaction Rate Subgrid-Scale Modeling

Victor Xing <sup>1,\*</sup> , Corentin Lapeyre <sup>1</sup> , Thomas Jaravel <sup>1</sup>  and Thierry Poinso <sup>1,2</sup> 

<sup>1</sup> CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse, France; lapeyre@cerfacs.fr (C.L.); jaravel@cerfacs.fr (T.J.); poinso@cerfacs.fr (T.P.)

<sup>2</sup> IMFT, Allée du Professeur Camille Soula, 31400 Toulouse, France

\* Correspondence: xing@cerfacs.fr

**Abstract:** Deep learning has recently emerged as a successful approach to produce accurate subgrid-scale (SGS) models for Large Eddy Simulations (LES) in combustion. However, the ability of these models to generalize to configurations far from their training distribution is still mainly unexplored, thus impeding their application to practical configurations. In this work, a convolutional neural network (CNN) model for the progress-variable SGS variance field is trained on a canonical premixed turbulent flame and evaluated *a priori* on a significantly more complex slot burner jet flame. Despite the extensive differences between the two configurations, the CNN generalizes well and outperforms existing algebraic models. Conditions for this successful generalization are discussed, including the effect of the filter size and flame–turbulence interaction parameters. The CNN is then integrated into an analytical reaction rate closure relying on a single-step chemical source term formulation and a presumed beta PDF (probability density function) approach. The proposed closure is able to accurately recover filtered reaction rate values on both training and generalization flames.

**Keywords:** large eddy simulation; turbulent combustion; deep learning; convolutional neural network; progress variable variance; generalization



**Citation:** Xing, V.; Lapeyre, C.; Jaravel, T.; Poinso, T. Generalization Capability of Convolutional Neural Networks for Progress Variable Variance and Reaction Rate Subgrid-Scale Modeling. *Energies* **2021**, *14*, 5096. <https://doi.org/10.3390/en14165096>

Academic Editor: Pinaki Pal

Received: 10 July 2021

Accepted: 10 August 2021

Published: 18 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In LES of turbulent premixed combustion, the large-scale disparity between the domain size and the flame thickness typically does not allow adequate resolution of the internal flame front structure. As a consequence, one of the main challenges for LES is the modeling of the SGS reaction source term. To this end, recent studies have successfully used machine learning to train artificial neural networks as SGS models. They have been applied to the modeling of scalar dissipation rates [1], filtered density functions [2,3], wrinkling functions [4,5], and have also been used to predict LES filtered reaction rates directly [6,7] or by deconvolution [8]. These studies have consistently shown that neural networks can outperform physical algebraic models on test cases identical or similar to their training configuration.

A major unknown for all machine learning models is their ability to generalize to unseen configurations that can greatly deviate from the data they were trained on. Existing works in combustion operate in a supervised learning framework [9], where Direct Numerical Simulations (DNS) are used to produce paired input/output data for the training procedure. If machine learning models cannot guarantee some form of generalizability, their predictive ability will be limited to applications for which DNS data is available, which typically require petascale levels of computing power for realistic configurations [10].

This paper addresses this issue by exploring under what conditions a model trained on a canonical turbulent flame performs on a very different complex configuration. A CNN model is built for the SGS variance of the progress variable, a key quantity in many

turbulent combustion closures [11–14]. Nikolaou et al. [8] had previously developed a CNN model for the variance, but it relied on the deconvolution of filtered quantities instead of direct computation of the variance, and was tested on cases similar to the training configuration.

The potential of the CNN model is illustrated by its integration in a new reaction rate closure. In a simplified single-step chemistry context, the mean LES reaction rate is analytically computed by a combination of a reaction source term formulation recently proposed by Pfitzner et al. [15,16] and a presumed beta probability density function (PDF) model, forming the so-called Pfitzner beta PDF CNN (PB-CNN) reaction rate closure. The CNN guarantees accurate predictions of the progress-variable variance, which parametrizes the beta PDF. Crucially, the analytical nature of the closure is expected to induce very fast evaluations, making this formulation a compelling alternative to tabulated chemistry methods that incur storage and computational costs [17–19].

The paper is organized as follows: the Pfitzner formulation for the reaction rate of a premixed flame is described in Section 2.1 while its coupling with a beta PDF model is presented in Section 2.2. Existing algebraic models for the progress variable variance are discussed in Section 2.3 before the rationale for using a CNN is discussed in Section 2.4. The training and generalization configurations are described in Section 3. The training procedure of the CNN is then detailed in Section 4. Results of the *a priori* evaluation of the CNN model and the PB-CNN closure are shown in Sections 5 and 6.

## 2. Combustion Modeling Framework

### 2.1. Pfitzner Source Term

The derivation of the chemical source term formulation proposed by Pfitzner et al. [15,16] is summarized here. Assuming adiabaticity, constant pressure, unity Lewis number, and single-step chemistry, the evolution of a steady, one-dimensional (1D) fully premixed flame is fully determined by a progress variable  $c = 1 - \frac{Y_F}{Y_F^f}$ , where  $Y_F$  is the fuel mass fraction and  $Y_F^f$  is its value in the fresh gases. The transport equation for  $c$  reads [11]

$$\rho u \frac{\partial c}{\partial x} = \frac{\partial}{\partial x} \left( \rho D \frac{\partial c}{\partial x} \right) - \frac{\dot{\omega}_F}{Y_F^f}, \quad (1)$$

where  $x, \rho, u, D, \dot{\omega}_F$  are the spatial coordinate, density, velocity, diffusivity, and fuel reaction rate, respectively. It is convenient to introduce a reduced spatial variable

$$\xi(x) = \int_0^x \frac{\rho_f s_L}{\rho(y) D(y)} dy, \quad (2)$$

with  $\rho_f$  the fresh gas density and  $s_L$  the laminar flame speed, in order to convert Equation (1) to a nondimensional form

$$\frac{\partial c}{\partial \xi} = \frac{\partial^2 c}{\partial \xi^2} - \frac{\rho D}{(\rho_f s_L)^2 Y_F^f} \dot{\omega}_F = \frac{\partial^2 c}{\partial \xi^2} + \omega(c), \quad (3)$$

where the nondimensional source term  $\omega(c)$  has been introduced. Pfitzner et al. [15,16] showed that  $c_m(\xi) = [1 + e^{-m\xi}]^{-1/m}$  is a solution of Equation (3) parametrized by a model coefficient  $m$ . It is associated with the nondimensional source term

$$\omega_m(c) = (m + 1)(1 - c^m)c^{m+1}, \quad (4)$$

which is a polynomial alternative to the exponential Arrhenius formulation typically used to compute the reaction rate. A constant prefactor is defined as  $\Lambda = (\rho_f s_L)^2 Y_F^f$  and the reaction rate  $\dot{\omega}_F$  is linked to the nondimensional source term  $\omega_m$  by

$$\dot{\omega}_F = -\frac{\Lambda}{\rho D} \omega_m. \quad (5)$$

Equation (5) describes what will be referred to as the Pfitzner reaction rate term. Pfitzner et al. [15,16] chose the model parameter  $m$  to match the profile of an Arrhenius reaction rate. For all  $m$  values, the laminar flame speed obtained from the total fuel consumption rate is equal to the input  $s_L$  value used in the prefactor  $\Lambda$ :

$$-\frac{1}{\rho_f Y_F^f} \int_{-\infty}^{\infty} \dot{\omega}_{F,m}(x) dx = s_L. \quad (6)$$

On the other hand, the thermal flame thickness  $\delta_L^0$  is inversely proportional to  $m$  and has the following analytical expression in  $\xi$  space

$$\delta_L^0 = \frac{(m+1)^{\frac{m+1}{m}}}{m}. \quad (7)$$

The Pfitzner formulation can therefore be a substitute to any flame computed with Arrhenius chemistry, by matching independently the flame speed with the  $\Lambda$  parameter and the flame thickness with the  $m$  parameter.

## 2.2. Pfitzner Beta PDF Closure for the Filtered Reaction Rate

In LES, any quantity of interest  $\phi$  is split into the sum of a mean and fluctuating component:  $\phi = \bar{\phi} + \phi'$ . Mean components are resolved by the numerical simulation on the LES mesh, while fluctuating components correspond to unresolved SGS activity that requires modeling. Mean components are related to the original quantity through a spatial filtering operation

$$\bar{\phi}(x) = \int_V G(x-x') \phi(x') dx', \quad (8)$$

where  $G$  is a low-pass filter with a cutoff scale equal to the LES mesh size. The filtered balance equations for reacting flows naturally introduce Favre-filtered (density-weighted) quantities  $\tilde{\phi} = \overline{\rho\phi}/\bar{\rho}$ .

The evolution of all species mass fractions can be recovered from the transport equation of the Favre-filtered fuel mass fraction  $\tilde{Y}_F$

$$\frac{\partial \tilde{\rho} \tilde{Y}_F}{\partial t} + \nabla \cdot (\tilde{\rho} \tilde{u} \tilde{Y}_F) = \nabla \cdot (\tilde{\rho} D \nabla \tilde{Y}_F) - \nabla \cdot (\tilde{\rho} \tilde{u} \tilde{Y}_F - \tilde{\rho} \tilde{u}' \tilde{Y}_F) + \tilde{\omega}_F, \quad (9)$$

where  $\tilde{\omega}_F$  is the filtered fuel reaction rate, which needs to be modeled to account for the unresolved SGS scales. To this end, PDF models rely on modeling the SGS distribution of  $c$ , noted as  $p(c)$ , to express the filtered reaction rate as the expected value of the DNS reaction rate

$$\tilde{\omega}_F = \int_0^1 p(c) \dot{\omega}_F(c) dc. \quad (10)$$

A common approach [20–24] is to presume that  $p(c)$  belongs to the family of beta PDFs that have the expression

$$p_\beta(c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} c^{a-1} (1-c)^{b-1}, \quad (11)$$

where  $\Gamma$  is the gamma function, and  $a$  and  $b$  are two parameters linked to the mean  $\bar{c}$  and variance  $\overline{c'^2} = \overline{c^2} - \bar{c}^2$  of the distribution by the relations

$$a = \bar{c} \left[ \frac{\bar{c}(1 - \bar{c})}{\overline{c'^2}} - 1 \right], \quad b = (1 - \bar{c}) \left[ \frac{\bar{c}(1 - \bar{c})}{\overline{c'^2}} - 1 \right]. \quad (12)$$

In combination with the Pfitzner source term, the beta PDF approach is particularly interesting because the LES nondimensional reaction rate  $(\omega_m)_\beta$  has an analytical expression [16]

$$\overline{(\omega_m)_\beta} = \int_0^1 p_\beta(c) \omega_m(c) \quad (13)$$

$$= (m + 1) \frac{\Gamma(a + b)}{\Gamma(a)} \left( \frac{\Gamma(a + m + 1)}{\Gamma(a + b + m + 1)} - \frac{\Gamma(a + 2m + 1)}{\Gamma(a + b + 2m + 1)} \right). \quad (14)$$

This is not the case with Arrhenius chemistry, and this property paves the way for computationally efficient evaluations of the filtered reaction rate. Using Equation (5), a similar formula can be derived for the fuel reaction rate under the simplifying assumption  $\rho D \approx \bar{\rho} \bar{D} = \bar{\rho} \hat{D}$ :

$$\overline{\dot{\omega}_F} = \int_0^1 p_\beta(c) \dot{\omega}_F \, dc = \Lambda \int_0^1 p_\beta(c) \frac{\omega_m(c)}{\rho(c) D(c)} \, dc \quad (15)$$

$$= \frac{\Lambda}{\bar{\rho} \bar{D}} \overline{(\omega_m)_\beta}. \quad (16)$$

with a Pfitzner beta PDF formulation, the turbulent LES reaction rate can be computed without needing on-the-fly integration of the PDF or tabulated chemistry.

### 2.3. Algebraic SGS Variance Models

To fully close the reaction rate, an SGS model is required for  $\overline{c'^2}$ . Many closures rely on algebraic expressions based on the gradient of the resolved progress variable. One such model [21,25,26] is given by

$$\overline{c'^2} = C_s \Delta^2 |\nabla \bar{c}|^2. \quad (17)$$

where  $\Delta$  is the local filter size and  $C_s$  is the model parameter. It is important to note that this model was originally proposed for nonreactive scalars, and for reactive scalars such as the progress variable, it should be modified to account for the correlation of the scalar with its source term [27]. Nevertheless, it is still commonly found in the literature whenever an algebraic model for the progress variable is needed [25,26,28–31]. It is therefore used here as a basis of comparison for the CNN, which plays the same role as a traditional algebraic model. The simplest way of estimating  $C_s$  is to use a global constant, and values such as 1/12 [26,28], 0.18 [25], or 0.5 [30] have been proposed in the literature. However, a constant value is not expected to be accurate for all flow configurations and filter sizes [26,29], and the best model constant is *a priori* unknown. The model of Equation (17) with a constant  $C_s$  parameter will be called the CST model.

$C_s$  can also be determined by a dynamic procedure [26,27,32,33] recalled in the following. A test filter of size  $\hat{\Delta} = 2\Delta$  is used to generate test-filtered quantities, noted as  $\hat{\phi}$ . It is assumed that the model coefficient varies slowly in space, and that Equation (17) holds at the filter and test filter levels with the same model coefficient. Its expression at the standard filter level is filtered to the test-filter level

$$\widehat{\overline{c'^2}} - \widehat{\bar{c}\bar{c}} = C_s \Delta^2 |\widehat{\nabla \bar{c}}|^2, \quad (18)$$

and independently expressed directly at the test-filter level

$$\widehat{\overline{c'^2}} - \widehat{\bar{c}\bar{c}} = C_s \hat{\Delta}^2 |\nabla \hat{c}|^2. \quad (19)$$

After introducing a Leonard term defined as

$$\mathcal{L} = \widehat{\bar{c}}\bar{c} - \widehat{\hat{c}}\hat{c}, \quad (20)$$

and a model term

$$\mathcal{M} = \hat{\Delta}^2 |\nabla \hat{c}|^2 - \Delta^2 |\nabla \bar{c}|^2, \quad (21)$$

the combination of Equations (18) and (19) leads to

$$\mathcal{L} = C_s \mathcal{M}. \quad (22)$$

As in [34], a least-squares procedure consisting of averaging over homogeneous directions of the flow is used to remedy stability issues. The final expression of the model coefficient is

$$C_s = \frac{\langle \mathcal{L} \mathcal{M} \rangle}{\langle \mathcal{M}^2 \rangle}, \quad (23)$$

where the averaging operation is denoted by the brackets  $\langle \cdot \rangle$ . The resulting model for  $\overline{c'^2}$  is called DYN.

#### 2.4. Convolutional Neural Networks for SGS Variance Modeling

The simple algebraic model formulation of Equation (17) is driven by sound physical arguments, but also limits the family of  $\overline{c'^2}$  fields that these models can fit. As a consequence, algebraic models are often inaccurate [26,31,33] and transport equations for  $\overline{c'^2}$  are usually preferred [35–37] but they induce additional computations and require supplementary modeling. To combine computational efficiency and accuracy,  $\overline{c'^2}$  is modeled using a CNN, which can be trained to model nearly arbitrarily complex functions due to its large representational capacity [9]. Compared with fully-connected architectures used in some studies for SGS modeling [3,7], CNNs benefit from locality and translation invariance inductive biases that are helpful to process grid elements [38], and have been empirically observed to generalize better on reconstruction tasks in computer vision [39].

A CNN model for  $\overline{c'^2}$  is trained on a planar flame wrinkled by homogeneous isotropic turbulence with simplified Pfitzner chemistry and evaluated *a priori* on the DNS of a slot burner jet flame with skeletal chemistry. Since most thermophysical and chemical parameters are purposefully different from one configuration to the other, this is a difficult generalization test for the model. It is assumed that local values of  $\overline{c'^2}$  can be determined from three main elements:

1. The profile of the  $\bar{c}$  field in the neighboring flame brush;
2. The amount of unresolved SGS scales;
3. The effect of local turbulence on the SGS distribution of  $c$ .

If they are known by the model or kept similar in the training and generalization configuration, the model should be able to generalize properly. The essential nature of these elements is highlighted by their presence in the algebraic model of Equation (17), respectively, in the form of the terms  $|\nabla \bar{c}|$ ,  $\Delta$ , and  $C_s$ . They therefore guide some of the key choices in the present work:

1. A CNN is used because of its ability to accurately learn spatial patterns in an extended area around a location of interest (Section 4.1);
2. the flame fronts seen in the training and evaluation contexts belong to the same turbulent combustion regime (Sections 3.1 and 3.2);
3. DNS snapshots are filtered and downsampled to a coarse grid with a well-chosen resolution (Section 4.2).

### 3. Training and Generalization Flow Configurations

#### 3.1. Training Configuration: Planar Flame in Homogeneous Isotropic Turbulence

The training database is built from the DNS of a freely-propagating, initially planar flame immersed in decaying homogeneous isotropic turbulence (HIT), similar to [31,40–42]. The cubic domain is uniformly meshed with  $384^3$  hexahedral elements and a mesh size  $\Delta x = 36 \mu\text{m}$ . In the  $x$  direction, NSCBC inlet and outlet boundary conditions are prescribed [43], while the other boundary conditions are periodic (Figure 1). The flame front is initialized along the  $x$  direction using a DNS solution of a 1D laminar propane–air flame at stoichiometry, temperature  $T = 300 \text{ K}$ , and pressure  $P = 1 \text{ bar}$ . Chemistry is represented by a single irreversible reaction  $\text{C}_3\text{H}_8 + 5 \text{O}_2 \rightarrow 3 \text{CO}_2 + 4 \text{H}_2\text{O}$  and the fuel reaction rate is computed using the Pfitzner DNS formulation of Equation (5). As key inputs needed in the formulation, the laminar flame speed  $s_L = 0.383 \text{ m s}^{-1}$  and the model parameter  $m = 3.8$  are chosen to match the laminar flame speed and flame thickness  $\delta_L = 352 \mu\text{m}$  obtained with single-step Arrhenius chemistry [44]. The flame thickness is based on the gradient of  $c$ :

$$\delta_L = \frac{1}{\max \frac{dc}{dx}}. \quad (24)$$

The initial laminar flame front in the HIT is therefore resolved on approximately 10 mesh points.

In the fresh gases, turbulent velocity fluctuations are superimposed on the velocity field of the laminar flame. They are generated by a von Karman–Pao spectrum [45]:

$$E(k) = \alpha \frac{u'^2}{k_e} \frac{(k/k_e)^4}{[1 + (k/k_e)^2]^{17/6}} \exp \left[ -2 \left( \frac{k}{k_\eta} \right)^2 \right], \quad (25)$$

where  $E(k)$ ,  $u'$ ,  $k_e$ ,  $k_\eta$  are the turbulent kinetic energy at the wavenumber  $k$ , turbulent fluctuation, wavenumber of the most energetic eddies, and wavenumber of the Kolmogorov scale  $\eta$ , respectively.  $\alpha = 1.453$  is a numerical constant. The integral length scale  $l_t$  is related to  $k_e$  through  $l_t \approx 0.747/k_e$  [45].  $u'$  and  $k_e$  are chosen so that

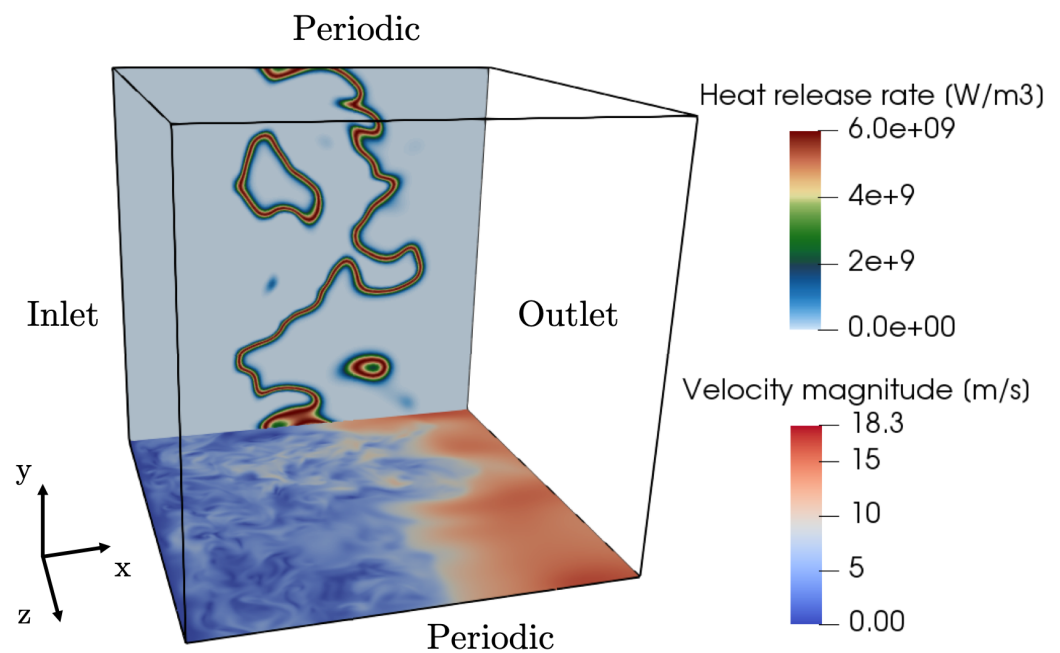
$$u'/s_L = 10, \quad l_t/\delta_L = 4.8, \quad (26)$$

leading to  $u' = 3.83 \text{ m s}^{-1}$  and  $k_e = 423 \text{ m}^{-1}$ .  $k_\eta$  is obtained via the dissipation rate  $\epsilon \approx u'^3/l_t$ , which leads to  $k_\eta = 2\pi/\eta = 2\pi\nu^{-3/4}\epsilon^{1/4} = 3.49 \times 10^5 \text{ m}^{-1}$ . The ratio of the mesh size to the Kolmogorov scale is  $\Delta x/\eta = 2$ , so that the mesh is fine enough to resolve all the eddies of the initial spectrum. Table 1 summarizes the main flame and turbulence parameters at initialization, including the Damköhler number  $Da = (l_t/\delta_L)/(u'/s_L)$ , Karlovitz number  $Ka = (u'/s_L)^{3/2}(l_t/\delta_L)^{-1/2}$ , turbulent Reynolds number  $Re_t = u'l_t/\nu$ , Taylor-scale Reynolds number  $Re_\lambda = u'\lambda_g/\nu$  based on the Taylor microscale  $\lambda_g$ , and eddy turnover time  $\tau = l_t/u'$ .

**Table 1.** HIT initial parameters.

$s_L$	$\delta_L$	$u'$	$l_t$	$\eta$	$Da$	$Ka$	$Re_t$	$Re_\lambda$	$\tau$
$0.383 \text{ m s}^{-1}$	$352 \mu\text{m}$	$3.83 \text{ m s}^{-1}$	$1.77 \text{ mm}$	$18 \mu\text{m}$	0.48	14	450	82	0.46 ms

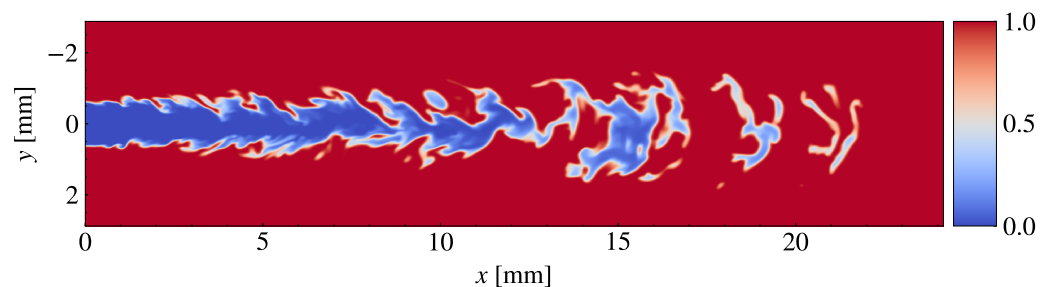
The DNS is run using the fully compressible explicit code AVBP [46,47] with the TTGC Taylor–Galerkin finite element scheme of third-order accuracy in time and space [48]. The simulation is run for 1.84 ms, corresponding to 4-eddy turnover times.



**Figure 1.** Slices of the HIT DNS at  $t = 2\tau$ : velocity magnitude in the  $xz$ -plane and heat release rate in the  $xy$ -plane.

### 3.2. Generalization Configuration: R2 Slot Burner Jet Flame

The configuration used to assess the generalizability of the CNN is the DNS of the R2 slot jet flame studied by Luca et al. [49–51]. A fully premixed methane–air mixture at temperature  $T = 800$  K, pressure  $P = 4$  bar, equivalence ratio  $\phi = 0.7$ , and bulk velocity  $U = 100 \text{ m s}^{-1}$  is injected through a slot of width  $H = 1.2$  mm and surrounded by a coflow of burnt gases. Chemistry is described by a skeletal chemical mechanism containing 16 species and 72 reactions [50]. The corresponding laminar flame has a speed of  $s_L = 1 \text{ m s}^{-1}$  and a thermal thickness  $\delta_L = 110 \text{ m s}^{-1}$ . The domain dimensions are 28.8 mm, 19.2 mm, and 5.16 mm in the streamwise ( $x$ ), crosswise ( $y$ ), and spanwise ( $z$ ) directions, respectively. It is uniformly meshed with a resolution  $\Delta x = 20 \text{ }\mu\text{m}$ . The progress-variable field of the fully developed jet flame is shown in Figure 2. 5 uncorrelated snapshots from instants where the jet is fully developed are used in this study, and form what will be called the *generalization set*.



**Figure 2.** Slice of progress-variable field in R2 DNS.

### 3.3. Comments on the Differences and Similarities between the Two Configurations

The HIT and R2 flames feature very different domain geometries, flow dynamics, and simulation parameters (Table 2). In addition, turbulence in the flame brush is induced by homogeneous isotropic turbulence in one case and shear between the jet and the coflow in the other. Previous studies involving CNNs investigated minor parametric variations in the inlet condition [4], fuel species, and Karlovitz number [1], or turbulence intensity [5,8]. Sub-



stantial generalization was observed by Wan et al. [7] for a fully connected neural network trained as a surrogate model for the filtered reaction rate on a micromixing database.

The present work investigates the capacity of the CNN trained on the HIT to generalize well to flames in the same premixed turbulent combustion regime, characterized by the velocity and length ratios  $u'/s_L, l_t/\delta_L$  [11,52,53]. To this end, in Equation (26), the parameters of the initial turbulence spectrum of the HIT were chosen to match the  $u'/s_L, l_t/\delta_L$  values computed by Luca et al. [51] at the crosswise location of maximum heat release and 60% of the flame length. They are considered to be representative of the turbulent combustion regime of the R2 flame.

**Table 2.** Main differences in the simulation parameters of the HIT and R2 DNS.

	HIT	R2
Fuel	$C_3H_8$	$CH_4$
Reactions	1	72
Species	5	16
$\phi$	1	0.7
$T$	300 K	800 K
$P$	1 bar	4 bar
$s_L$	$0.383 \text{ m s}^{-1}$	$1 \text{ m s}^{-1}$
$\delta_L$	$352 \mu\text{m}$	$85 \mu\text{m}$
$\Delta x$	$36 \mu\text{m}$	$20 \mu\text{m}$

Finally, three orders of magnitude separate the computational cost of the two DNS simulations. To simulate 0.9 ms of physical time, R2 requires 1 million CPU hours on an Intel Xeon Haswell-based supercomputer [49], while the HIT only needs 1600 CPU hours on a comparable Intel Xeon Skylake-based cluster. This highlights the greatly reduced computational demands of training an ML model on simple canonical configurations instead of full-scale realistic flames.

## 4. Machine Learning Framework

### 4.1. U-Net Architecture

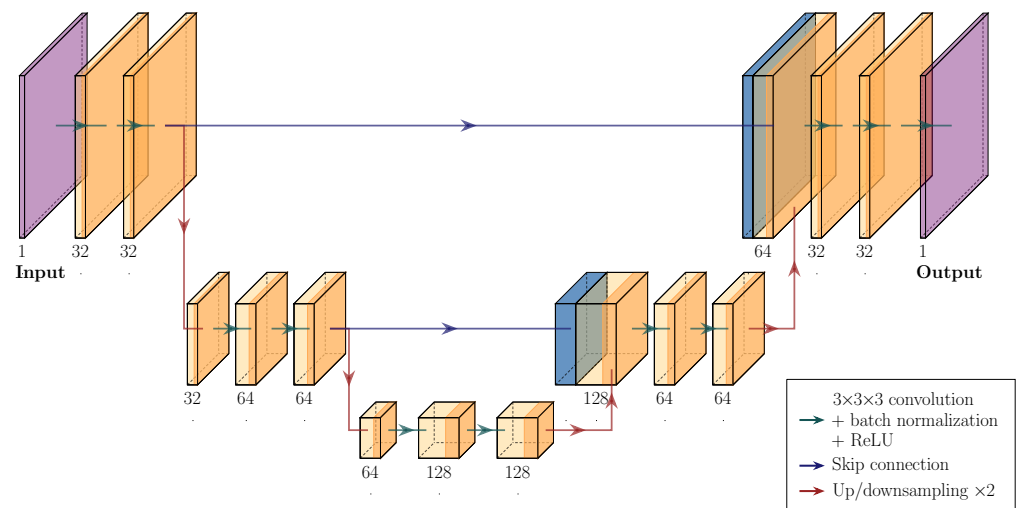
The CNN trained in this study is a U-Net [54] with the same architecture used by Lapeyre et al. [4] to model SGS flame wrinkling. U-Nets are convolutional neural networks that are well-suited to field-to-field visual pattern recognition tasks, and are commonly used in computer vision for 2D and 3D image segmentation [54–56]. They rely on a succession of convolutional layers and downsampling operations to learn visual patterns in a wide area around the location of the prediction [4,9]. This is an appealing approach for combustion SGS modeling since the CNN is able to make local predictions based on the knowledge of the full spatial structure of the flame front.

The full architecture is detailed in Figure 3. It follows a fully convolutional, symmetrical encoder–decoder structure. Each stage is composed of two successions of

- A 3D convolution with a  $3 \times 3 \times 3$  kernel with “same” padding;
- A batch normalization layer [57];
- A rectified linear unit (ReLU) nonlinear activation unit;
- A  $2 \times 2 \times 2$  max-pooling downsampling operation in the encoder or a  $2 \times 2 \times 2$  upsampling operation in the decoder.

The specificity of the U-Net architecture lies in the skip connections that concatenate the end of each encoder stage to the start of the corresponding decoder stage. Due to the downsampling operations, each stage of the encoder processes feature maps at decreasing resolutions. The combination of the upsampling and concatenation encourages each decoder stage to aggregate information from two separate spatial scales. Skip connections also allow the network to combine low-level features learned by the shallow layers of the encoder with the more complex, abstract features learned by the decoder, and

accelerate training convergence [56]. Multiscale CNN architectures for field-to-field fluid predictions were successfully used by Lapeyre et al. [4] to model SGS flame wrinkling and Ajuria et al. [58] to solve the Poisson equation in incompressible flows. In the following, the U-Net model will simply be called the CNN. In the context of SGS modeling, this multiscale approach can be thought of as an extension of the dynamic procedure, which extracts information at a test-filtered scale to inform local predictions. Its ability to predict the entire output field at once is also attractive in terms of computational efficiency, as a single inference of the network is needed for the whole domain.



**Figure 3.** U-Net architecture. Each feature map is represented with 2 out of its 3 spatial dimensions as height and depth, and the channel dimension as width. The number of channels is noted below each feature map. The hidden layers connecting the feature maps are represented by the arrows. The input and output fields are colored in pink, and the feature maps concatenated by the skip connections are colored in blue. The network contains a total of 1.4 million trainable parameters.

#### 4.2. Data Preparation

The CNN is trained to process  $\bar{c}$  input fields and predict associated  $\overline{c^2}$  output fields. Following the LES framework, the filtered fields are represented on a grid of size  $\Delta$  that is coarser than the DNS mesh. To generate these filtered coarse fields, 3D fields of  $c$  are first extracted from instantaneous snapshots of the DNS simulations.  $c$  is computed as

$$c_{HIT} = 1 - \frac{Y_F}{Y_F^f}, \quad (27)$$

with the fresh gas fuel mass fraction  $Y_F^f$ .

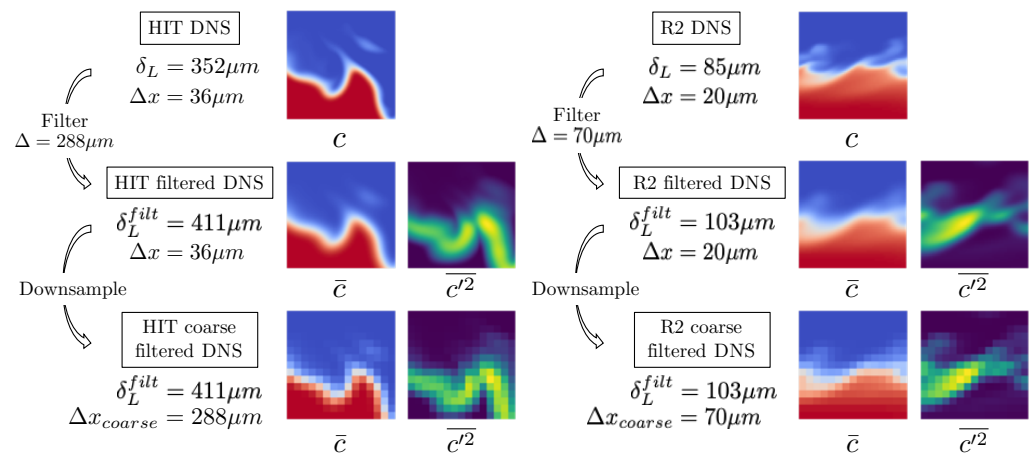
The DNS progress variable fields are then filtered with a Gaussian filter

$$G(x) = \left( \frac{6}{\pi\Delta^2} \right)^{3/2} \exp \left[ \frac{-6x^2}{\Delta^2} \right], \quad (28)$$

and downsampled to a coarse-mesh-structured grid with a resolution  $\Delta$ .

The filtering and coarsening procedure for HIT and R2 is illustrated in Figure 4. The value of the coarse mesh resolution for the HIT and R2 is chosen to ensure that the ratio  $\Delta/\delta_L = 0.81$  is the same for both cases. This is key to the ability of the CNN to generalize from one configuration to the other despite their different flame thickness and DNS mesh sizes. Since  $\Delta/\delta_L$  is kept constant, the ratio of filtered laminar flame thickness to filter size  $\delta_L^{filt}/\Delta$  is also approximately equal. It can be interpreted as the number of coarse mesh points on which the filtered flame brushes are resolved. This therefore guarantees that similar flame front patterns and gradients are found in the coarse HIT and R2, allowing

the CNN to transfer from one case to the other. Additionally, the conservation of  $\Delta/\delta_L$  establishes that these similar flame front gradients are unequivocally related to equivalent SGS activity, and equivalent values of  $\overline{c'^2}$  predicted by the CNN. In this regard, the filter sizes are chosen as equal to  $288\ \mu\text{m}$  and  $70\ \mu\text{m}$  for the HIT and R2, respectively.



**Figure 4.** Summary of the filtering process for the HIT and R2 for  $\Delta/\delta_L = 0.81$ , illustrated with enlargements of the corresponding flame.

#### 4.3. Training Procedure

The training database is built from 46 instantaneous snapshots of the HIT DNS, extracted at regular intervals of 0.04 ms. The first 38 constitute the *training* set for the CNN, while the next 4 form the validation set and the final 4 are kept as a hold-out *test* set for the results shown below. Instead of directly predicting  $\overline{c'^2}$ , the CNN is trained to predict the unmixedness factor  $\overline{c'^2}/(\overline{c}(1 - \overline{c}))$ , a normalized version of the variance that was empirically found to lead to slightly better results. As a preprocessing step, the input and output fields are periodically padded in the  $y$  and  $z$  directions to enforce the periodicity boundary conditions of the DNS.

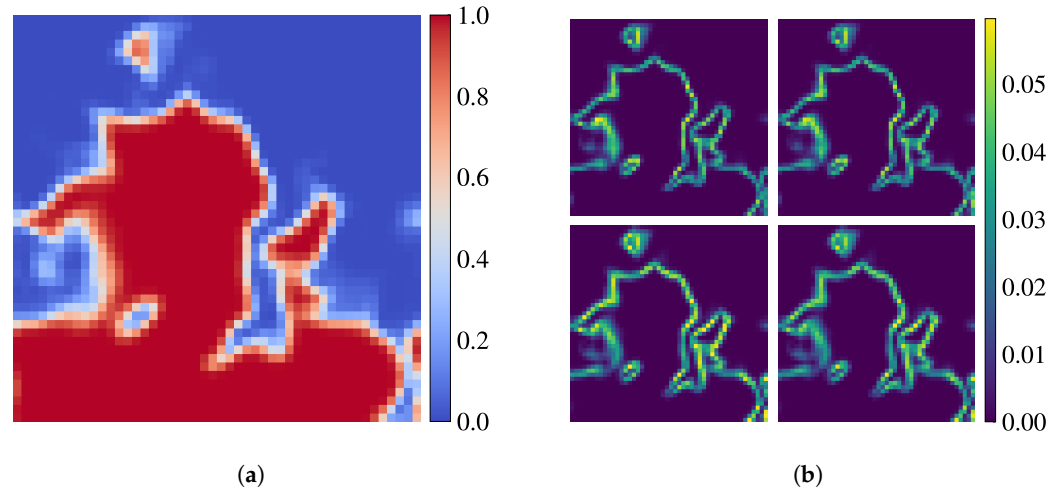
At each training iteration, 16 snapshots from the training set are sampled and, for each snapshot, 4 randomly cropped  $32^3$  cubes are passed through random  $90^\circ$  rotations and mirror operations before being added to the training batch. Gradient descent optimization is performed by an Adam optimizer [59] with an initial learning rate of 0.01. A mean squared error loss function encourages the CNN outputs to match the ground truth  $\overline{c'^2}$  fields. Training was performed with the Tensorflow 2 Python library and stopped after 10,000 iterations when convergence was well-established. On an NVIDIA Tesla V100 GPU, this equates to one hour of training time. This training procedure only needs to be done once, after which the weights of the CNN are frozen and the trained model can predict  $\overline{c'^2}$  from a given  $\overline{c}$  field in a few milliseconds.

## 5. A Priori Evaluation of the $\overline{c'^2}$ Model

### 5.1. Evaluation on the HIT Test Set

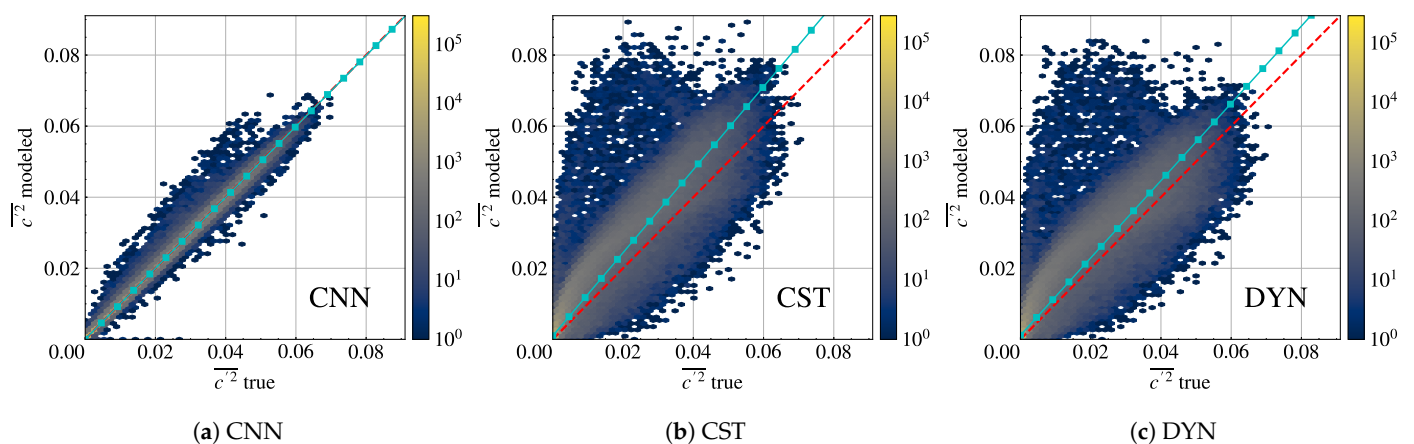
The *a priori* performance of the CNN on the test set of the HIT is assessed to ensure that it has successfully learned to represent  $\overline{c'^2}$  on samples from the same distribution as the training database. The CNN model predictions are compared to ground truth  $\overline{c'^2}$  values computed from the filtered downsampled DNS, called *true* values, and to the CST model predictions. The constant coefficient  $C_s = 0.18$  [25] was chosen by picking the value in the literature [25,26,28,30] that best matched the true  $\overline{c'^2}$  values. For the DYN model, since there are no fully homogeneous directions, the averaging procedure is conducted on isocontours of  $\overline{c}$ . Another possibility is to perform averaging over the whole model, leading to a dynamically determined constant  $C_s = 0.17$ , which is very close to the value chosen for the CST model.

Figure 5 shows sample slices of  $\bar{c}$ , true  $\overline{c'^2}$ , and modeled  $\overline{c'^2}$  fields from the test set. The CNN matches the ground truth faithfully in all regions of the flow, while the gradient model does not seem as accurate.



**Figure 5.** Sample slices of  $\bar{c}$  (a), and  $\overline{c'^2}$  (b): ground truth (top-left), CNN (top-right), CST (bottom-left), and DYN (bottom-right) model on the HIT test set.

Model predictions are aggregated on the whole test set and plotted against true values in Figure 6. The excellent performance of the CNN is confirmed by the tight spread of its predictions around the perfect model line, and the lack of bias of their linear trend. In comparison, CST and DYN model predictions display significantly higher variance and nonzero bias. Note that with prior knowledge of the ground truth, a  $C_s$  value that leads to an unbiased CST model can be chosen but this would have little effect on the variance. The dynamic procedure proves to be effective at estimating a local model coefficient without any prior knowledge and leads to a slightly smaller error than the CST model.



**Figure 6.** Hexbin plots of model predictions against ground truth: (a) CNN model, (b) CST model, (c) DYN model. Data points are gathered in hexagonal bins and colored according to the number of data points they contain. The red line  $y = x$  represents a perfect model. Cyan lines represent linear trends of the model.

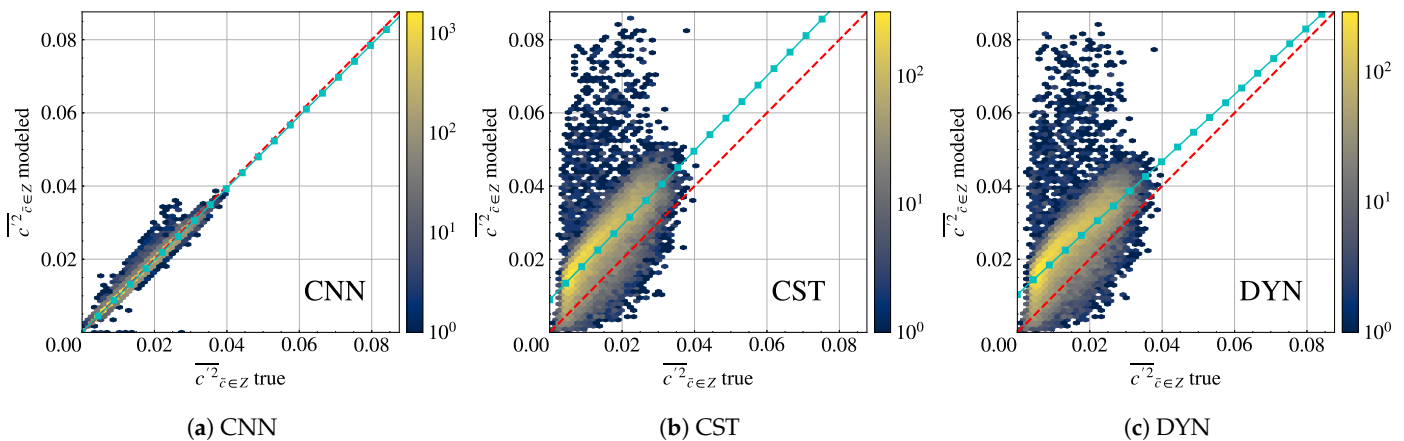
Most of the data points plotted in Figure 6 are in regions of the flow where no flame front is present and using  $\overline{c'^2}$  for SGS closure is not important. The critical zones for  $\overline{c'^2}$  models are located in the reaction zone of the flame, identified as the regions where the laminar Pfitzner reaction rate is greater than 10% of its maximum value:

$$\mathcal{Z} = \{\bar{c} : \dot{\omega}_F(\bar{c}) \geq 0.1 \times \max_{0 \leq c \leq 1} (\dot{\omega}_F(c))\}. \quad (29)$$

Hexplots conditioned on  $\bar{c} \in \mathcal{Z}$  are plotted in Figure 7, and error values are also reported in Table 3. It is interesting to note that the CNN performs even better in these critical regions, and is very close to a perfect model. On the other hand, the performances of the gradient models are significantly worse. Error values computed using a normalized mean square metric

$$\epsilon = \frac{\sum (y_{\text{model}} - y_{\text{true}})^2}{\sum y_{\text{true}}^2} \quad (30)$$

are reported in Table 3 and confirm the excellent *a priori* performance of the CNN on the HIT test set.



**Figure 7.** Hexbin plots of model predictions against ground truth conditioned on  $\bar{c} \in \mathcal{Z}$ : (a) CNN model, (b) CST model, (c) DYN model.

**Table 3.** HIT test set error values for  $\bar{c}^2$ .

Model	$\epsilon$	$\epsilon, \bar{c} \in \mathcal{Z}$
CNN	0.060	0.054
CST	0.417	0.733
DYN	0.398	0.730

### 5.2. Evaluation on the R2 Generalization Set

The various models are now assessed on the 5 snapshots of R2 composing the generalization set. For the DYN model, averaging is performed in the spanwise direction of the burner.

Figure 8 shows a sample slice of the  $\tilde{c}$  field in a filtered snapshot of R2 and the predictions of the CNN, CST, and DYN model for the corresponding SGS variance. Once again, the value  $C_s = 0.18$  is retained for the CST model. All models reproduce qualitatively the evolution of the filtered DNS  $\bar{c}^2$ . The hexplots of Figure 9 and the error values reported in Table 4 reveal that the CNN performs the best out of all the models, which is a demonstration of its ability to transfer its predictive power from the HIT flame to the R2 flame. However, it is naturally less accurate than on the test of the HIT, which contained samples from the same distribution as its training set. In comparison, the CST and DYN models show zero bias but a higher variance. Unlike in the HIT, all the models perform similarly when evaluated on the whole domain or only the reaction zone  $\mathcal{Z}$ , as indicated by the error values in Table 4.

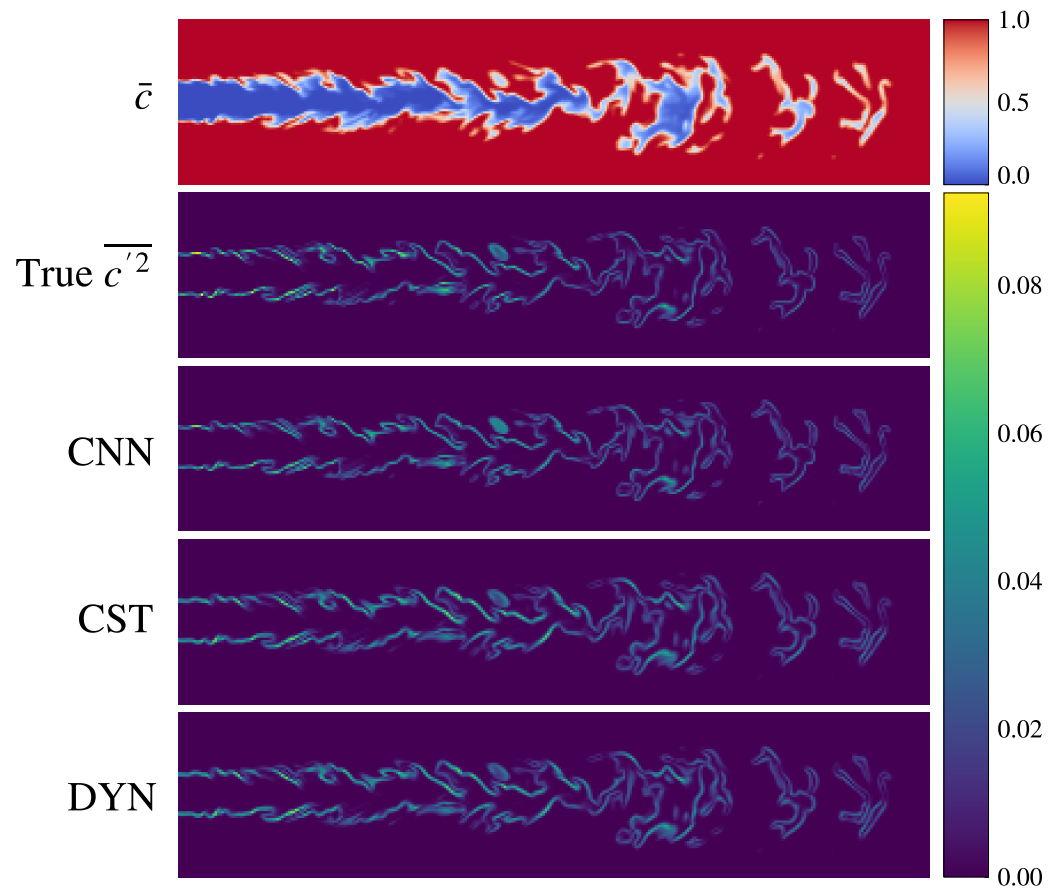


Figure 8. Sample R2 slices of  $\bar{c}$ , true  $c^2$ , CNN, CST, and DYN model predictions.

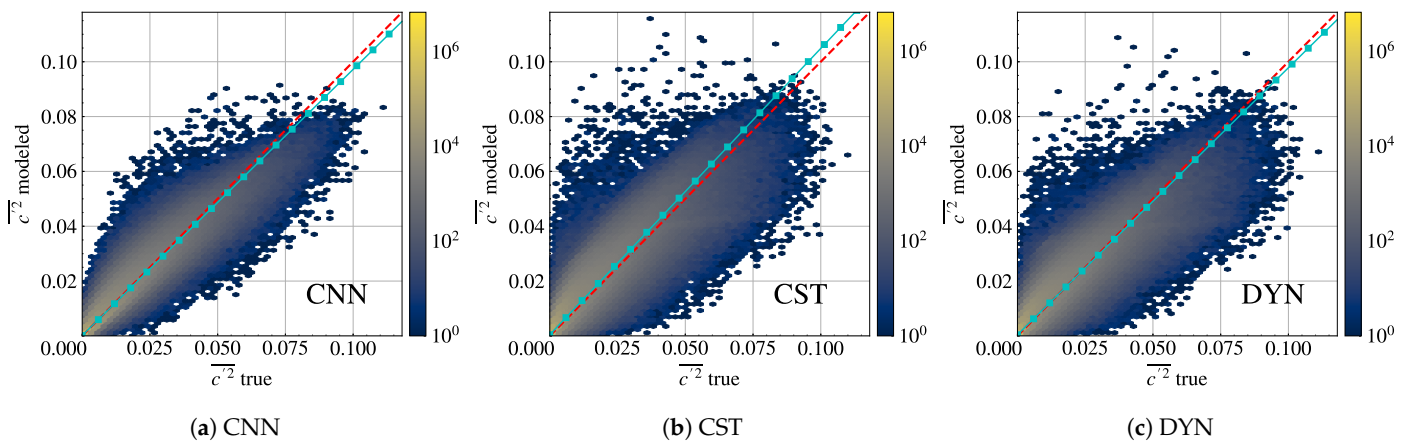


Figure 9. Hexbin plots of model predictions against ground truth on the R2 generalization set: (a) CNN model, (b) CST model, (c) DYN model.

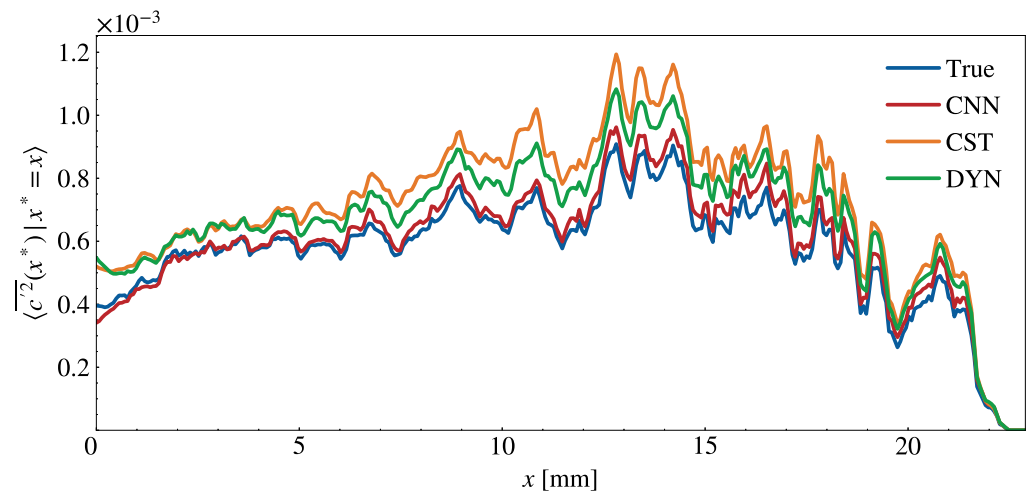
Table 4. R2 error values for  $c^2$ .

Model	$\epsilon$	$\epsilon, \bar{c} \in \mathcal{Z}$
CNN	0.241	0.214
CST	0.371	0.331
DYN	0.358	0.322

For a sample snapshot, the spatial distribution of model predictions is investigated by plotting  $\bar{c}^2$  values averaged on transverse slices in Figure 10. The CNN recovers the

filtered DNS values nearly perfectly, while the algebraic models consistently overestimate the variance.

This *a priori* evaluation of  $\overline{c'^2}$  models has shown that the CNN trained on the HIT configuration was able to generalize accurately to the R2 flame and outperform constant coefficient and dynamic algebraic models.



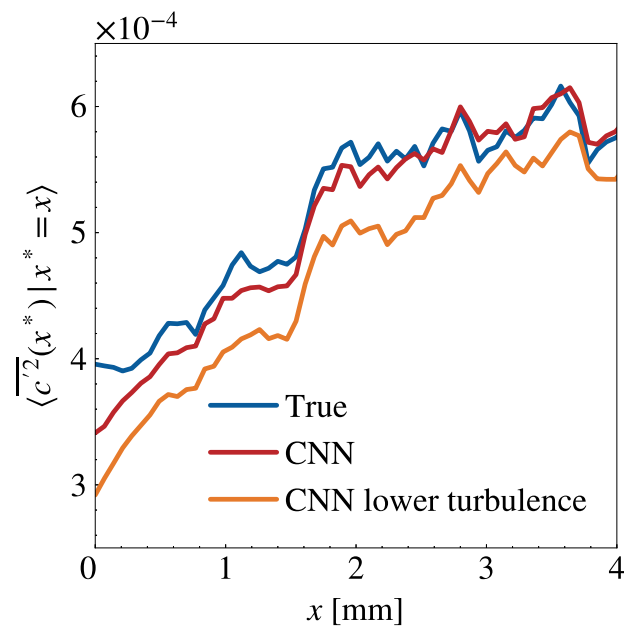
**Figure 10.** Evolution of transverse averages of the true  $\overline{c'^2}$  and model predictions along the streamwise direction  $x$ .

### 5.3. Discussion on the Conditions for Generalization

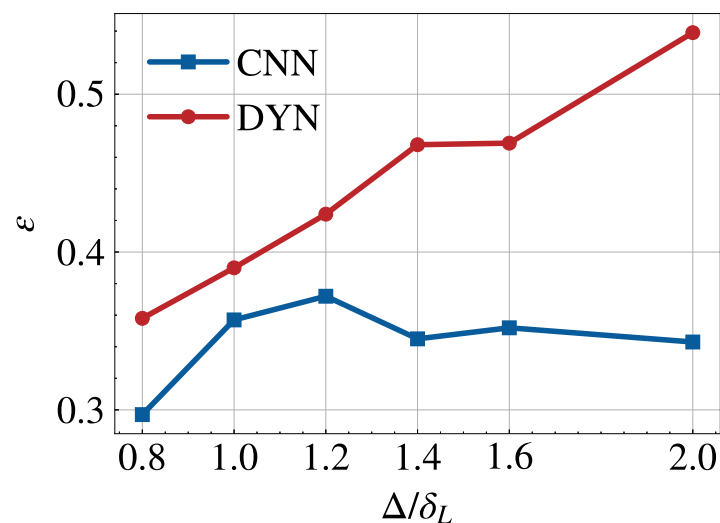
Between the HIT training/test sets and the R2 generalization set, three key ratios are conserved:  $u'/s_L$ ,  $l_t/\delta_L$ , and  $\Delta/\delta_L$ . It is important to understand the sensitivity of the generalization of the CNN to these ratios, as they could place strict limits on the practical applicability of the model.

The influence of  $u'/s_L$  and  $l_t/\delta_L$  is analyzed by training a second CNN on the same HIT configuration with halved initial values of  $u'/s_L$ ,  $l_t/\delta_L$ , leading to lesser wrinkling in the turbulent flame front. When evaluated on R2, this second CNN had an error  $\epsilon = 0.301$ , which is 30% higher than the reference CNN. Figure 11 shows that the effect of training on a weaker HIT is noticeable in the first 4 mm of the flame near the inlet. The second CNN underpredicts the SGS variance compared with the reference CNN and the filtered DNS. This is coherent with the fact that it is trained on a weaker HIT, which contains lower  $\overline{c'^2}$  values than the original configuration. This brief study seems to indicate that the choice of  $u'/s_L$ ,  $l_t/\delta_L$  is indeed impactful on the generalization accuracy of the CNN, but more extensive work should be performed to fully understand the effect of these ratios.

The restriction to relying on a single value of  $\Delta/\delta_L$  can be relaxed by training the CNN on a range of filter size values. An instance of the CNN was trained on a dataset comprising the collection of training HIT snapshots, which were filtered at  $\Delta/\delta_L = 0.8$ , 1.2 and 1.6, for a total of 126 snapshots with 3 separate filter sizes. It was then tested on R2 for the same filter sizes, as well as unseen values  $\Delta/\delta_L = 1$ , 1.4, and 2. Figure 12 shows the evolution of the error made by the CNN and DYN models with  $\Delta/\delta_L$ . When  $\Delta/\delta_L$  increases, the performance of the DYN model deteriorates as the SGS modeling task becomes more difficult. In contrast, the error made by the CNN is stable across all filter sizes, including unseen ones. This suggests that with proper training, the CNN can be accurate on a range of  $\Delta/\delta_L$  values instead of a single one. Note that the error made by this CNN on  $\Delta/\delta_L = 0.8$  is higher than the one reported in Table 4 for the CNN solely trained on this filter size. This implies that a balance must be found between accuracy at a single filter size, and validity across a range of values. Interestingly, providing the  $\Delta/\delta_L$  value as an additional input channel for the CNN did not improve the results. The rest of this work is presented with the original CNN trained with  $\Delta/\delta_L = 0.8$ .



**Figure 11.** Evolution of transverse averages of the true  $\overline{c^2}$  and both CNN predictions in the first 4 mm of the R2 flame.



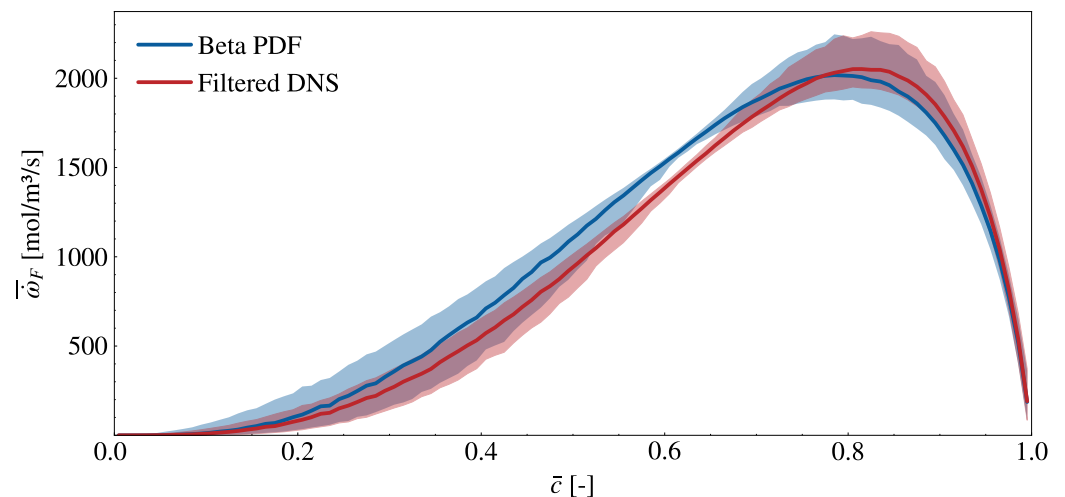
**Figure 12.** CNN and DYN model errors on R2 for various filter sizes.

## 6. A Priori Evaluation of the PB-CNN Model for $\overline{\omega}_F$

### 6.1. Evaluation on the HIT Test Set

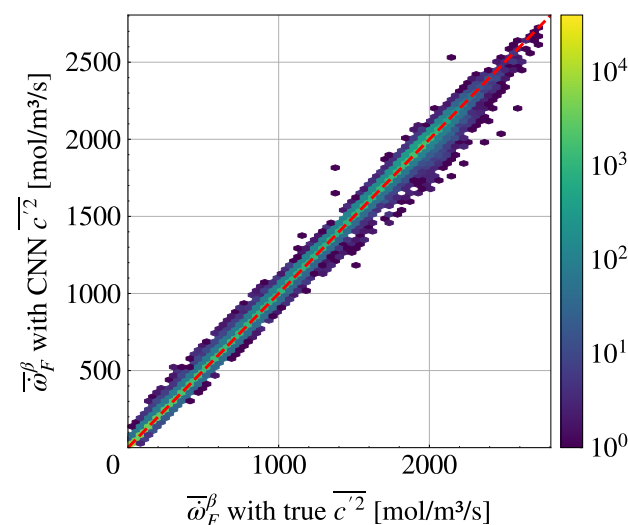
In the previous section, the CNN was shown to be able to learn a model for  $\overline{c^2}$ , which is accurate on the test and generalization configurations. In the following, modeled values of  $\overline{c^2}$  are incorporated in the presumed beta PDF approach detailed in Section 2.2 to form the PB-CNN model for  $\overline{\omega}_F$ . Beta PDFs are known to sometimes be inaccurate models for the SGS distribution of  $c$  in premixed combustion [22,26,35]. Therefore, the presumed beta PDF approach needs to be justified independently from the model for  $\overline{c^2}$ . For the HIT, Figure 13 shows the conditional median, 5% and 95% quantiles plotted against  $\bar{c}$  for the true filtered reaction rate, and the Pfitzner beta PDF reaction rate with true  $\overline{c^2}$  values. The beta PDF appears to be a good approximation for the true SGS PDF, showing that the presumed beta PDF assumption is sound in the case of the HIT.





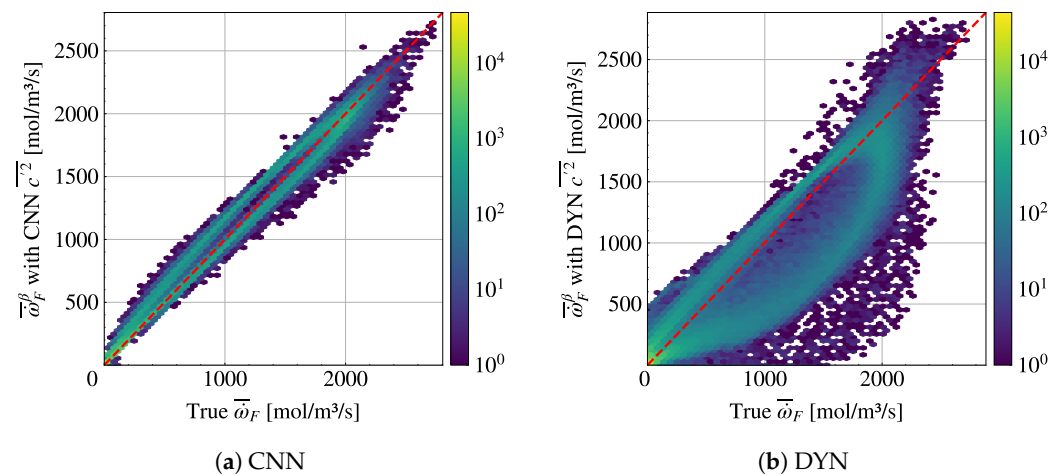
**Figure 13.** Distribution of the filtered reaction rate computed from a beta PDF assumption with true  $\bar{c}^{T2}$  (blue) and the filtered DNS (red): conditional mean (solid line); 5% and 95% conditional quantiles (delimiting the shaded area). The statistics are aggregated over the HIT test set.

Given the accuracy of the CNN model, replacing the true  $\bar{c}^{T2}$  values with the CNN predictions does not noticeably affect the shape of the beta PDF. The isolated effect of the CNN model on the beta PDF reaction rate appears minimal in Figure 14, where the beta PDF reaction rates with CNN values of  $\bar{c}^{T2}$  are compared with beta PDF reaction rates with true values of  $\bar{c}^{T2}$ .

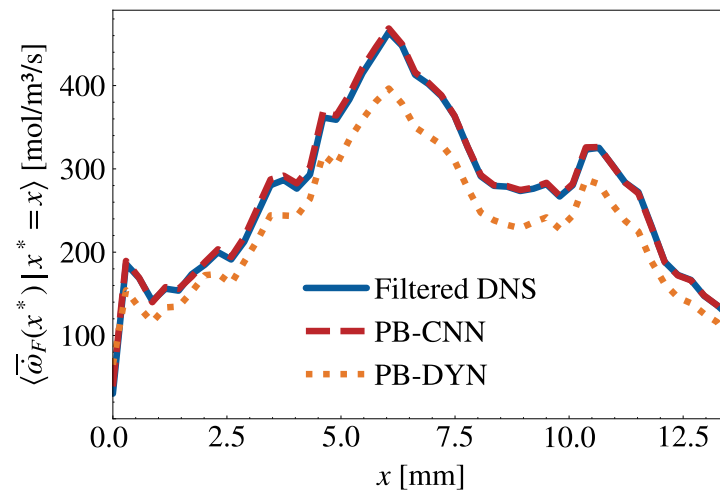


**Figure 14.** Hexbin plot of beta PDF reaction rates with CNN values of  $\bar{c}^{T2}$  against beta PDF reaction rates with true values of  $\bar{c}^{T2}$ .

Finally, the combined effect of the beta PDF assumption and the  $\bar{c}^{T2}$  model is shown in Figure 15. The high variance values predicted by the DYN model lead to underpredictions of the reaction rate when combined with the Pfitzner beta PDF closure. In contrast, using CNN predictions leads to a much closer fit to the true filtered reaction rate. Figure 16, where true and modeled reaction rates averaged on  $yz$ -planes are plotted against the streamwise  $x$  coordinate, illustrates how the quality of the CNN model for  $\bar{c}^{T2}$  leads to a good very estimate of fuel consumption rates using a beta PDF assumption.



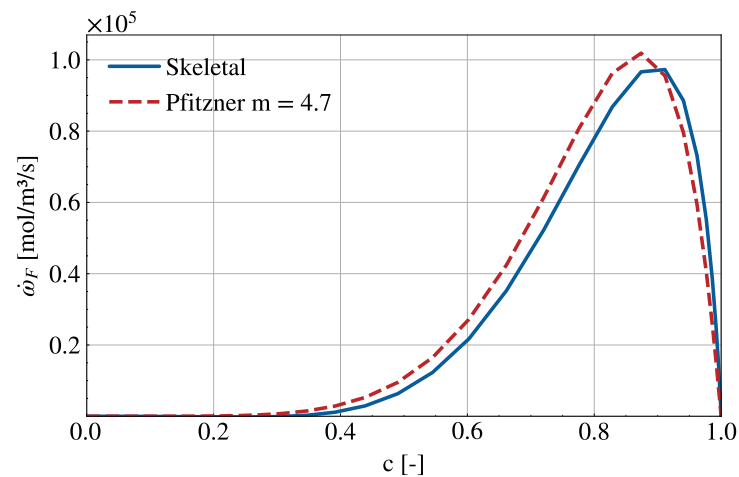
**Figure 15.** Hexbin plots comparing true filtered reaction rates to beta PDF reaction rates using modeled  $\overline{c^2}$  values: (a) CNN model, (b) DYN model.



**Figure 16.** Evolution in the streamwise  $x$  direction of the true and modeled reaction rates averaged on  $yz$ -planes, for a sample in the HIT test set. PB-DYN denotes the Pfitzner beta PDF closure with the DYN model for  $\overline{c^2}$ .

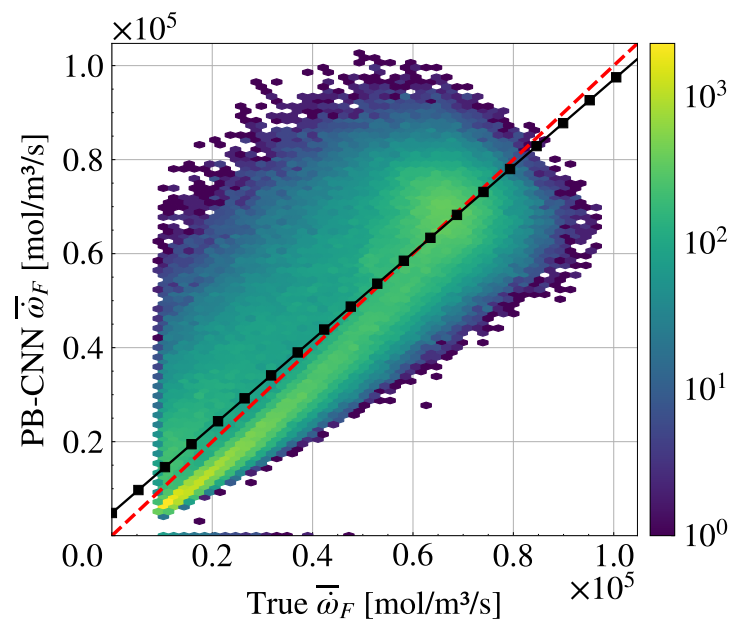
## 6.2. Evaluation on the R2 Generalization Set

The PB-CNN closure is finally evaluated on R2. From the R2 coarse fields of  $\overline{c}$ , beta PDF reaction rates are computed according to Equation (15) and using either true  $\overline{c^2}$  or predictions from the CNN. A reference 1D laminar flame matching the conditions of the R2 simulation is used to compute the prefactor  $\Lambda = (\rho_{fSL})^2 Y_F^f$ . The value for the Pfitzner model parameter  $m = 4.7$  was chosen so that the total fuel consumption rate of the reference 1D flame is matched by the Pfitzner reaction rate based on the  $c$  field of the reference 1D flame. As a result, by construction, both reaction rates would lead to the same total fuel consumption rate on a planar laminar flame. Figure 17 shows the close match of the resulting Pfitzner reaction rate to the fuel reaction rate of the skeletal mechanism.



**Figure 17.** DNS fuel reaction rates for skeletal and Pfitzner chemistry vs. progress variable in a 1D laminar flame.

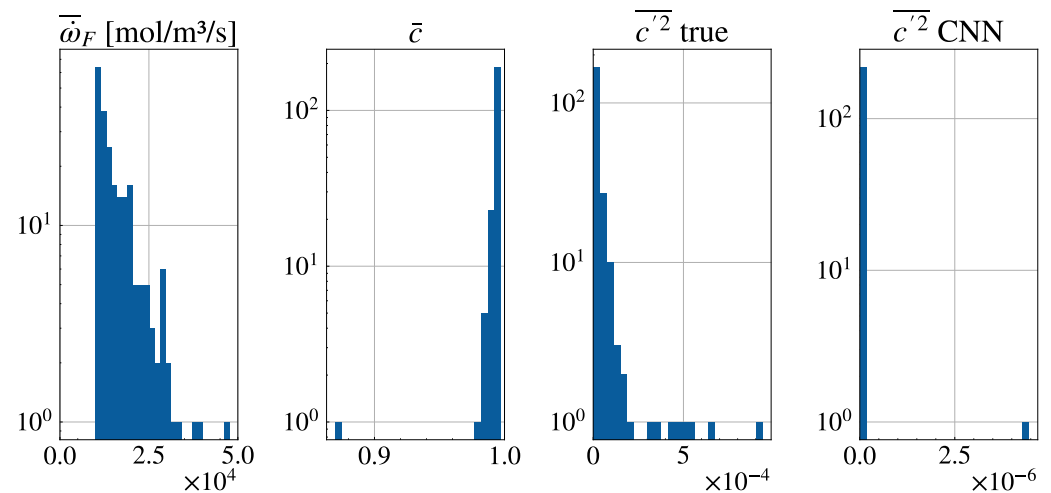
Figure 18 compares the PB-CNN reaction rates with the true filtered values of R2. For clarity, only the points where the filtered DNS reaction rate is greater than 10% of its maximum value are plotted. This corresponds to the same condition as Equation (29). Compared with the HIT flame, the scatter plot still exhibits low bias, with a large portion of the predictions concentrated around the reference linear trend, but it also has a higher variance, which is emphasized by the logarithmic scale of the color bar. This indicates that some local values of the reaction rate are not well predicted by the PB-CNN model.



**Figure 18.** Hexbin plots comparing true filtered reaction rates with PB-CNN reaction rates for points in the reaction zone of the filtered DNS of R2. The black line represents the linear trend of the model.

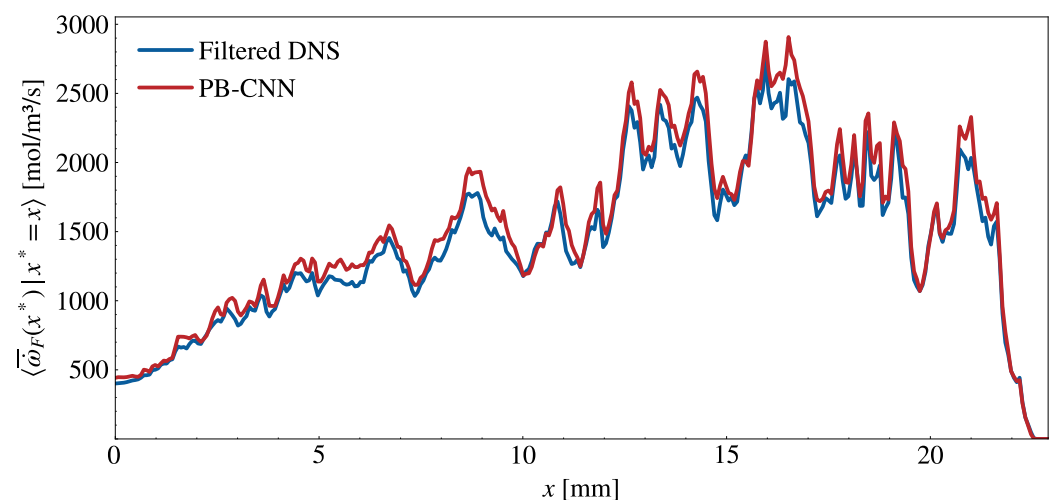
One particular group of mispredicted points is located on the x-axis of the plot and corresponds to regions in the reaction zone where the PB-CNN model erroneously predicts a reaction rate of exactly 0. Analyzing the reason for this misprediction is insightful to understand the behavior of the model in difficult edge cases. This group of points is characterized in Figure 19 by the distribution of their values of the true filtered reaction rate, filtered progress variable, true SGS variance, and CNN predictions for the variance. It appears that these points are located in the postflame region, where  $c \approx 1$  and the skeletal mechanism still predicts some chemical activity. On the other hand, the SGS variance is

extremely small and estimated by the CNN to be 0, causing the beta PDF-based reaction rate to also be exactly 0. This highlights a limitation of comparing a simplified reaction rate formulation relying on the SGS variance against a skeletal mechanism. However, these mispredictions are a rare occurrence since this group of points only constitutes 0.1% of the points in the reaction zone.



**Figure 19.** Histograms of  $\overline{\omega}_F$ ,  $\bar{c}$ , true  $\overline{c'^2}$ , and CNN  $\overline{c'^2}$  for the points in the reaction zone where the PB-CNN reaction rate is exactly zero.

Despite discrepancies in local predictions of the reaction rate, the PB-CNN model provides correct estimates of spatially averaged reaction rates. This is important as it shows that the model would be useful in LES to recover key integral quantities of the burner such as the mean length of the flame. This is shown in Figure 20, which compares the streamwise evolution of the filtered DNS and PB-CNN reaction rates averaged on transverse slices. The PB-CNN closure recovers the correct averaged reaction rate in all regions of the flame. This leads to a fuel consumption rate in the whole domain within 5% of the reference value. *A posteriori* simulations could be carried out in the future to assess if the PB-CNN reaction rate leads to an LES flame with the same length and total burning rate as the R2 DNS.



**Figure 20.** Evolution of transverse averages of skeletal and beta PDF reaction rates along the streamwise direction  $x$ .

## 7. Conclusions

A deep convolutional neural network was trained to model the progress variable SGS variance  $\overline{c'^2}$  on a planar flame in homogeneous isotropic turbulence, and was shown

to generalize accurately to the R2 complex slot burner jet flame in an *a priori* evaluation. Despite very different flow dynamics, thermophysical parameters, and chemistry, the conservation of the normalized filter size  $\Delta/\delta_L$  and the turbulence–flame interaction ratios  $u'/s_L, l_t/\delta_L$  were assumed to be key to the generalizability of the CNN, and their influence on the results was discussed. This work indicates that CNNs can be trained on canonical simple cases and used in practical configurations. Future studies could focus on applying this methodology to different combustion models. The universal nature of the conditions for generalization that are exhibited here should also be investigated, as well as ways to integrate them into the machine learning model.

In an effort to combine computational efficiency and modeling accuracy, the PB-CNN closure for the mean LES fuel reaction rate was proposed. It combines the reaction source term formulation of Pfitzner et al. [15,16], a presumed beta PDF assumption, and the CNN model for  $\overline{c^2}$  in an analytical formula for the reaction rate. It showed excellent *a priori* results on the R2 flame when compared with the burning rate produced by the original skeletal chemistry. In the future, an *a posteriori* evaluation would be a challenging test of the accuracy of the  $\overline{c^2}$  model given by the CNN and the PB-CNN closure for  $\overline{\omega_F}$ . This formulation would be especially well suited to simulating large-scale premixed combustion configurations, such as explosions in venting chambers. The resulting overpressure can be accurately captured using simplified chemistry, but remains a challenge at scale due to insufficient mesh resolution, prohibitive computational costs, and strong sensitivity to turbulent reaction rates [44,60].

**Author Contributions:** Conceptualization, V.X., C.L., T.J., and T.P.; methodology, V.X., T.P.; software, V.X.; validation, V.X.; formal analysis, V.X.; investigation, V.X.; resources, V.X.; data curation, V.X.; writing—original draft preparation, V.X.; writing—review and editing, V.X., C.L., T.J., and T.P.; visualization, V.X.; supervision, C.L., T.J., and T.P.; project administration, T.P.; funding acquisition, V.X. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by an AMX fellowship of Ecole polytechnique.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** The authors gratefully acknowledge Antonio Attili for providing the DNS data of the R2 flame.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

CNN	Convolutional Neural Network
DNS	Direct Numerical Simulation
HIT	Homogeneous Isotropic Turbulence
LES	Large Eddy Simulation
PB-CNN	Pfitzner Beta PDF CNN
PDF	Probability Density Function
ReLU	Rectified Linear Unit
SGS	Subgrid-Scale

## References

1. Yellapantula, S.; Perry, B.A.; Grout, R.W. Deep learning-based model for progress variable dissipation rate in turbulent premixed flames. *Proc. Combust. Inst.* **2020**, *38*, 1–10. [CrossRef]
2. de Frahan, M.T.H.; Yellapantula, S.; King, R.; Day, M.S.; Grout, R.W. Deep learning for presumed probability density function models. *Combust. Flame* **2019**, *208*, 436–450. [CrossRef]
3. Chen, Z.X.; Iavarone, S.; Ghiasi, G.; Kannan, V.; D'Alessio, G.; Parente, A.; Swaminathan, N. Application of machine learning for filtered density function closure in MILD combustion. *Combust. Flame* **2021**, *225*, 160–179. [CrossRef]
4. Lapeyre, C.J.; Misdariis, A.; Cazard, N.; Veynante, D.; Poinot, T. Training convolutional neural networks to estimate turbulent sub-grid scale reaction rates. *Combust. Flame* **2019**, *203*, 255–264. [CrossRef]
5. Ren, J.; Wang, H.; Luo, K.; Fan, J. A priori assessment of convolutional neural network and algebraic models for flame surface density of high Karlovitz premixed flames. *Phys. Fluids* **2021**, *33*, 036111. [CrossRef]
6. Seltz, A.; Domingo, P.; Vervisch, L.; Nikolaou, Z.M. Direct mapping from LES resolved scales to filtered-flame generated manifolds using convolutional neural networks. *Combust. Flame* **2019**, *210*, 71–82. [CrossRef]
7. Wan, K.; Barnaud, C.; Vervisch, L.; Domingo, P. 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* **2020**, *220*, 119–129. [CrossRef]
8. Nikolaou, Z.M.; Chrysostomou, C.; Vervisch, L.; Cant, S. Progress Variable Variance and Filtered Rate Modelling Using Convolutional Neural Networks and Flamelet Methods. *Flow Turbul. Combust.* **2019**, *103*, 485–501. [CrossRef]
9. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
10. Chen, J.H. Petascale direct numerical simulation of turbulent combustion—Fundamental insights towards predictive models. *Proc. Combust. Inst.* **2011**, *33*, 99–123. [CrossRef]
11. Poinot, T.; Veynante, D. *Theoretical and Numerical Combustion*, 3rd ed.; 2011. Available online: [www.cerfacs.fr/elearning](http://www.cerfacs.fr/elearning) (accessed on 1 July 2021).
12. Moureau, V.; Domingo, P.; Vervisch, L. From Large-Eddy Simulation to Direct Numerical Simulation of a lean premixed swirl flame: Filtered laminar flame-PDF modeling. *Combust. Flame* **2011**, *158*, 1340–1357. [CrossRef]
13. Langella, I.; Swaminathan, N. Unstrained and strained flamelets for LES of premixed combustion. *Combust. Theory Model.* **2016**, *20*, 410–440. [CrossRef]
14. Klimenko, A.Y.; Bilger, R.W. Conditional moment closure for turbulent combustion. *Prog. Energy Combust. Sci.* **1999**, *25*, 595–687. [CrossRef]
15. Pfitzner, M.; Klein, M. A near-exact analytic solution of progress variable and pdf for single-step Arrhenius chemistry. *Combust. Flame* **2021**, *226*, 380–395. [CrossRef]
16. Pfitzner, M. A New Analytic pdf for Simulations of Premixed Turbulent Combustion. *Flow Turbul. Combust.* **2021**, *106*, 1213–1239. [CrossRef]
17. Pope, S.B. Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation. *Combust. Theory Model.* **1997**, *1*, 41–63. [CrossRef]
18. Fiorina, B.; Vicquelin, R.; Auzillon, P.; Darabiha, N.; Gicquel, O.; Veynante, D. A filtered tabulated chemistry model for LES of premixed combustion. *Combust. Flame* **2010**, *157*, 465–475. [CrossRef]
19. van Oijen, J.A.; Donini, A.; Bastiaans, R.J.; ten Thijs Boonkcamp, J.H.; de Goey, L.P. State-of-the-art in premixed combustion modeling using flamelet generated manifolds. *Prog. Energy Combust. Sci.* **2016**, *57*, 30–74. [CrossRef]
20. Bray, K.N.C. Studies of the turbulent burning velocity. *Proc. R. Soc. Lond. Ser. A Math. Phys. Sci.* **1990**, *431*, 315–335. [CrossRef]
21. Cook, A.W.; Riley, J.J. A subgrid model for equilibrium chemistry in turbulent flows. *Phys. Fluids* **1994**, *6*, 2868–2870. [CrossRef]
22. Bray, K.N.; Champion, M.; Libby, P.A.; Swaminathan, N. Finite rate chemistry and presumed PDF models for premixed turbulent combustion. *Combust. Flame* **2006**, *146*, 665–673. [CrossRef]
23. Proch, F.; Kempf, A.M. Numerical analysis of the Cambridge stratified flame series using artificial thickened flame LES with tabulated premixed flame chemistry. *Combust. Flame* **2014**, *161*, 2627–2646. [CrossRef]
24. Langella, I.; Swaminathan, N.; Williams, F.A.; Furukawa, J. Large-Eddy Simulation of Premixed Combustion in the Corrugated-Flamelet Regime. *Combust. Sci. Technol.* **2016**, *188*, 1565–1591. [CrossRef]
25. Veynante, D.; Knikker, R. Comparison between LES results and experimental data in reacting flows. *J. Turbul.* **2006**, *7*, 1–20. [CrossRef]
26. Knudsen, E.; Kim, S.H.; Pitsch, H. An analysis of premixed flamelet models for large eddy simulation of turbulent combustion. *Phys. Fluids* **2010**, *22*, 115109. [CrossRef]
27. Pierce, C.D.; Moin, P. A dynamic model for subgrid-scale variance and dissipation rate of a conserved scalar. *Phys. Fluids* **1998**, *10*, 3041–3044. [CrossRef]
28. Balarac, G.; Pitsch, H.; Raman, V. Development of a dynamic model for the subfilter scalar variance using the concept of optimal estimators. *Phys. Fluids* **2008**, *20*, 035114. [CrossRef]
29. Vreman, A.W.; Bastiaans, R.J.; Geurts, B.J. A similarity subgrid model for premixed turbulent combustion. *Flow Turbul. Combust.* **2009**, *82*, 233–248. [CrossRef]
30. Ranjan, R.; Muralidharan, B.; Nagaoka, Y.; Menon, S. Subgrid-Scale Modeling of Reaction-Diffusion and Scalar Transport in Turbulent Premixed Flames. *Combust. Sci. Technol.* **2016**, *188*, 1496–1537. [CrossRef]

31. Keil, F.B.; Klein, M.; Chakraborty, N. Sub-grid Reaction Progress Variable Variance Closure in Turbulent Premixed Flames. *Flow Turbul. Combust.* **2021**, *106*, 1195–1212. [[CrossRef](#)]
32. Wall, C.; Boersma, B.J.; Moin, P. An evaluation of the assumed beta probability density function subgrid-scale model for large eddy simulation of nonpremixed, turbulent combustion with heat release. *Phys. Fluids* **2000**, *12*, 2522–2529. [[CrossRef](#)]
33. Mukhopadhyay, S.; Van Oijen, J.A.; De Goeij, L.P. A comparative study of presumed PDFs for premixed turbulent combustion modeling based on progress variable and its variance. *Fuel* **2015**, *159*, 728–740. [[CrossRef](#)]
34. Lilly, D.K. A proposed modification of the Germano subgrid-scale closure method. *Phys. Fluids A Fluid Dyn.* **1992**, *4*, 633–635. [[CrossRef](#)]
35. Domingo, P.; Vervisch, L.; Payet, S.; Hauguel, R. DNS of a premixed turbulent V flame and LES of a ducted flame using a FSD-PDF subgrid scale closure with FPI-tabulated chemistry. *Combust. Flame* **2005**, *143*, 566–586. [[CrossRef](#)]
36. Domingo, P.; Vervisch, L.; Veynante, D. Large-eddy simulation of a lifted methane jet flame in a vitiated coflow. *Combust. Flame* **2008**, *152*, 415–432. [[CrossRef](#)]
37. Nilsson, T.; Langella, I.; Doan, N.A.K.; Swaminathan, N.; Yu, R.; Bai, X.S. A priori analysis of sub-grid variance of a reactive scalar using DNS data of high Ka flames. *Combust. Theory Model.* **2019**, *23*, 885–906. [[CrossRef](#)]
38. Battaglia, P.W.; Hamrick, J.B.; Bapst, V.; Sanchez-Gonzalez, A.; Zambaldi, V.; Malinowski, M.; Tacchetti, A.; Raposo, D.; Santoro, A.; Faulkner, R.; et al. Relational inductive biases, deep learning, and graph networks. *arXiv* **2018**, arXiv:1806.01261.
39. Zhang, C.; Bengio, S.; Hardt, M.; Mozer, M.C.; Singer, Y. Identity Crisis: Memorization and Generalization Under Extreme Overparameterization. In Proceedings of the International Conference on Learning Representations, Addis Ababa, Ethiopia, 26–30 April 2020.
40. Boger, M.; Veynante, D.; Boughanem, H.; Trouvé, A. Direct numerical simulation analysis of flame surface density concept for large eddy simulation of turbulent premixed combustion. *Symp. Int. Combust.* **1998**, *27*, 917–925. [[CrossRef](#)]
41. Charlette, F.; Meneveau, C.; Veynante, D. A power-law wrinkling model for LES of premixed turbulent combustion Part I: Non-dynamic formulation and initial tests. *Combust. Flame* **2002**, *131*, 159–180. [[CrossRef](#)]
42. Chakraborty, N.; Cant, R.S. Effects of Lewis number on scalar transport in turbulent premixed flames. *Phys. Fluids* **2009**, *21*, 1–12. [[CrossRef](#)]
43. Poinso, T.J.; Lele, S.K. Boundary conditions for direct simulations of compressible viscous flows. *J. Comput. Phys.* **1992**, *101*, 104–129. [[CrossRef](#)]
44. Quillatre, P.; Vermorel, O.; Poinso, T.; Ricoux, P. Large eddy simulation of vented deflagration. *Ind. Eng. Chem. Res.* **2013**, *52*, 11414–11423. [[CrossRef](#)]
45. Bailly, C.; Juvé, D. A stochastic approach to compute subsonic noise using linearized euler's equations. In Proceedings of the 5th AIAA/CEAS Aeroacoustics Conference and Exhibit, Reston, VA, USA, 10–12 May 1999; pp. 496–506. [[CrossRef](#)]
46. Schönfeld, T.; Rudgyard, M. Steady and unsteady flow simulations using the hybrid flow solver AVBP. *AIAA J.* **1999**, *37*, 1378–1385. [[CrossRef](#)]
47. Selle, L.; Lartigue, G.; Poinso, T.; Koch, R.; Schildmacher, K.U.; Krebs, W.; Prade, B.; Kaufmann, P.; Veynante, D. Compressible large eddy simulation of turbulent combustion in complex geometry on unstructured meshes. *Combust. Flame* **2004**, *137*, 489–505. [[CrossRef](#)]
48. Colin, O.; Rudgyard, M. Development of High-Order Taylor-Galerkin Schemes for LES. *J. Comput. Phys.* **2000**, *162*, 338–371. [[CrossRef](#)]
49. Luca, S.; Attili, A.; Bisetti, F. Direct Numerical Simulation of Turbulent Lean Methane-Air Bunsen Flames with Mixture Inhomogeneities. In Proceedings of the 54th AIAA Aerospace Sciences Meeting AIAA, San Diego, CA, USA, 4–8 January 2016. [[CrossRef](#)]
50. Luca, S.; Al-Khateeb, A.N.; Attili, A.; Bisetti, F. Comprehensive Validation of Skeletal Mechanism for Turbulent Premixed Methane–Air Flame Simulations. *J. Propuls. Power* **2017**, *34*, 153–160. [[CrossRef](#)]
51. Luca, S.; Attili, A.; Lo Schiavo, E.; Creta, F.; Bisetti, F. On the statistics of flame stretch in turbulent premixed jet flames in the thin reaction zone regime at varying Reynolds number. *Proc. Combust. Inst.* **2019**, *37*, 2451–2459. [[CrossRef](#)]
52. Peters, N. Laminar flamelet concepts in turbulent combustion. *Symp. Int. Combust.* **1988**, *21*, 1231–1250. [[CrossRef](#)]
53. Peters, N. The turbulent burning velocity for large-scale and small-scale turbulence. *J. Fluid Mech.* **1999**, *384*, 107–132. [[CrossRef](#)]
54. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9351, pp. 234–241. [[CrossRef](#)]
55. Çiçek, Ö.; Abdulkadir, A.; Lienkamp, S.S.; Brox, T.; Ronneberger, O. 3D U-net: Learning dense volumetric segmentation from sparse annotation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; pp. 424–432. [[CrossRef](#)]
56. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 3–11. [[CrossRef](#)]
57. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6–11 July 2015; Volume 1, pp. 448–456.

- 
58. Ajuria, E.; Alguacil, A.; Bauerheim, M.; Misdariis, A.; Cuenot, B.; Benazera, E. Towards a hybrid computational strategy based on deep learning for incompressible flows. In Proceedings of the AIAA AVIATION 2020 FORUM, Virtual, 15–19 June 2020; pp. 1–17. [[CrossRef](#)]
  59. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015-Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015; pp. 1–15.
  60. Vermorel, O.; Quillatre, P.; Poinso, T. LES of explosions in venting chamber: A test case for premixed turbulent combustion models. *Combust. Flame* **2017**, *183*, 207–223. [[CrossRef](#)]