Energy and AI 16 (2024) 100341

Contents lists available at ScienceDirect

Energy and AI

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

Utilizing neural networks to supplant chemical kinetics tabulation through mass conservation and weighting of species depletion

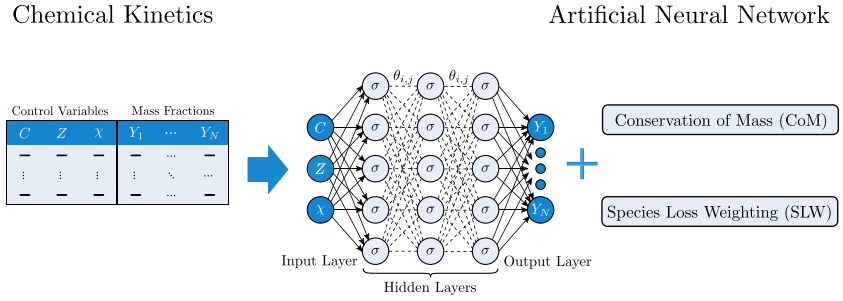
Franz M. Rohrhofer [a,*], Stefan Posch [b], Clemens Gößnitzer [b], José M. García-Oliver [c], Bernhard C. Geiger [a]

[a] Know-Center GmbH, Sandgasse 36, Graz, 8010, Styria, Austria
[b] LEC GmbH, Inffeldgasse 19, Graz, 8010, Styria, Austria
[c] Universitat Politècnica de València, Camí de Vera, València, 46022, València, Spain

HIGHLIGHTS

- A neural network approach for replacing tabulated chemical kinetics is studied.
- A mass-conserving architecture and species loss weighting are suggested.
- The performance of the proposed approach is validated on four reaction mechanisms.
- The approach shows an excellent performance on predicting major and minor species mass fractions.
- The approach also scales well to complex reaction mechanisms involving multiple species.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Neural network approach
Chemical kinetics
Flamelet tabulation
Mass conservation
Species loss weighting

ABSTRACT

Artificial Neural Networks (ANNs) have emerged as a powerful tool in combustion simulations to replace memory-intensive tabulation of integrated chemical kinetics. Complex reaction mechanisms, however, present a challenge for standard ANN approaches as modeling multiple species typically suffers from inaccurate predictions on minor species. This paper presents a novel ANN approach which can be applied on complex reaction mechanisms in tabular data form, and only involves training a single ANN for a complete reaction mechanism. The approach incorporates a network architecture that automatically conserves mass and employs a particular loss weighting based on species depletion. Both modifications are used to improve the overall ANN performance and individual prediction accuracies, especially for minor species mass fractions. To validate its effectiveness, the approach is compared to standard ANNs in terms of performance and ANN complexity. Four distinct reaction mechanisms (H_2, C_7H_{16}, $C_{12}H_{26}$, OME_{34}) are used as a test cases, and results demonstrate that considerable improvements can be achieved by applying both modifications.

1. Introduction

The application of artificial neural networks (ANNs) in chemical combustion systems exhibits significant potential for addressing the increasing computational demands associated with complex fluid-flow simulations [1]. With the ongoing advance in deep learning, complex reaction mechanisms can now be deployed in computational fluid dynamics (CFD) with reduced demands for CPU time and RAM utilization [2,3]. Among several approaches, ANNs have shown great applicability to various types of chemical systems, e.g., in the study

* Corresponding author.
E-mail address: frohrhofer@acm.org (F.M. Rohrhofer).

https://doi.org/10.1016/j.egyai.2024.100341
Received 21 September 2023; Received in revised form 6 December 2023; Accepted 22 January 2024
Available online 30 January 2024

Nomenclature

T	Temperature
Z	Mixture fraction
C	Progress variable
χ	Scalar dissipation rate
Y_k	Mass fraction of species k
$\boldsymbol{Y}$	Mass fractions
$\boldsymbol{X}$	Neural network input (features)
$\boldsymbol{Y}_\theta$	Neural network output (targets)
$\boldsymbol{\theta}$	Neural network parameters
L	Number of layers
ϕ	Activation function
$\mathcal{L}$	Loss function
ω	Loss weights
ρ	Pearson product-moment correlation coefficient (PPMCC)

of oxygenate additives [4] or hydrocarbon chemistry [5]. Furthermore, ANNs have also been successfully coupled with existing methods including design of experiments (DOE) [6,7] and tabulation methods which are in focus of this work. Tabulation methods utilize lookup tables that represent precomputed chemical kinetics through large databases which are retrieved many times during a simulation. Taking up several hundreds or thousands of Megabytes (MB), these tables drastically raise the memory utilization which can be a limiting factor for complex reaction mechanisms in distributed memory simulations. Several tabulation methods have been developed to reduce the computational demands, such as the Flamelet Generated Manifold (FGM) [8,9], Rate-Controlled Constrained Equilibrium (RCCE) [10,11], or Self-Organizing-Maps (SOM) [12,13]. More details on these methods can be found in [1]. Often, these tabulation methods are then combined with ANNs that are capable of learning and predicting the generated chemical kinetics, providing smooth interpolations with only a small number of network parameters needed. For example, in [14] an ANN technique is combined with the FGM method to simulate spray combustion using large eddy and Reynolds averaged Navier–Stokes simulations. Achieving results that agree very well with the experiments, the ANN approach only consumes a fraction of the memory utilization compared to the conventional FGM method. Also, in [15] a ANN model is combined with the FGM method to study water sprayed turbulent combustion, and similar conclusions are drown from the results. Furthermore, the incorporation of physical constraints in ANNs has been studied in [16], and it has been found to increase performances in predicting tabulated chemical kinetics.

Nevertheless, employing a single ANN for an entire reaction mechanism often presents a challenge in terms of ANN optimization and accuracy. Simultaneously learning from heterogeneous target scales commonly results in dominant targets being learned more accurately, while minor ones may be learned less effectively [2,17]. This has been reported multiple times in combustion studies and is frequently associated with the depletion of minor species in complex reaction mechanisms [18,19]. The use of several ANNs, also known as the multiple multilayer perceptrons or MMLP approach [20,21], has been demonstrated to mitigate this issue, by applying separate ANNs to either individual targets or clustered components of a reaction mechanism. This may increase the performance on specific targets, such as minor species mass fractions, as the ANN optimization is facilitated by focusing on either single targets or targets that share chemical kinetics. However, with the number of ANNs the complexity for training and model employment also increases which can be a limiting factor when considering reaction mechanisms with many species involved. The trade-off between computational complexity and accuracy has always been a prominent factor in chemical combustion systems, and it appears that ANNs are not exempt from this consideration.

In this paper, a novel ANN approach is proposed for replacing lookup tables which is simple and involves only a single ANN for a complete reaction mechanism. The ANN shares its network parameters across several species involved in the reaction mechanism, thus representing the underlying reaction mechanism with a single ANN. The approach further comprises a particular network architecture that accounts for conservation of mass (CoM) and species loss weighting (SLW) that is used in the ANN optimization procedure. The ANN performance is validated on four different lookup tables that are generated by the flamelet approach and represent four distinct reaction mechanisms, namely H_2, $C_{12}H_{26}$, OME_{34}, and C_7H_{16}. Furthermore, comparison to standard ANN approaches are provided in terms of accuracy and computational efficiency. In all experiments, the proposed ANN approach shows an excellent prediction accuracy on individual species and demonstrates great scalability to complex chemical systems.

2. Numerical methods

2.1. Flamelet model

The flamelet model, initially derived by Peters [22], assumes that a turbulent flame is an ensemble of laminar one-dimensional reacting structures, referred to as ‘flamelets’. Flamelets are solved externally in ad-hoc solvers, from which the thermochemical state in terms of temperature T and species mass fraction Y_k can be obtained along the flamelet evolution. This information can be used to describe laminar combustion processes, but it is commonly extended to create advanced turbulent-chemistry interaction models, where a CFD code interacts with the flamelet solver results to obtain the thermochemical state in the turbulent flame. This approach overcomes the need to solve a transport equation for each species in every cell of the computational domain.

Some flamelet methods are tabulated to allow for accelerated CFD-flamelet interactions during simulation run-time. The information exchange occurs over a small number of the so-called ’control variables’. Depending on the flame configuration, different approaches may be followed; for example, in the current approach an Unsteady Flamelet Progress Variable (UFPV) tabulation depending on mixture fraction Z (local fuel-air ratio), progress variable C (chemical evolution from inert to steady conditions) and scalar dissipation rate χ (related to the spatial gradient along the 1D flamelet domain) is followed [23–25].

2.2. Artificial neural network

Artificial neural networks (ANNs) are a versatile class of artificial intelligence methods and, with fuzzy logic and genetic algorithms, are most widely used in prediction and classification tasks cross diverse scientific disciplines [26,27]. ANNs usually consist of multiple layers and neurons (or perceptrons), with their interconnections specified by the ANN architecture. Common in the field of combustion chemistry are multilayer perceptrons (MLPs) that are a simple class of feedforward ANNs with multiple layers of neurons that are fully-connected, see Fig. 1. This network architecture is particularly well-suited for approximating nonlinear mappings $\boldsymbol{X} \mapsto \boldsymbol{Y}$ as given by lookup tables, where $\boldsymbol{X}$ is given by the input control variable (here C, Z and χ) and $\boldsymbol{Y}$ by the respective species mass fractions Y_k. The task of training the ANN is to find a good approximation to this nonlinear mapping by using training examples sampled from the table. The network function with L layers is given by the recursive application of activations

$$\boldsymbol{Y}_\theta = \phi^{(L)}(\boldsymbol{\theta}^{(L)}, \cdots, \phi^{(1)}(\boldsymbol{\theta}^{(1)}, \phi^{(0)}(\boldsymbol{\theta}^{(0)}, \boldsymbol{X})) \cdots), \tag{1}$$

where $\phi^{(l)}(\cdot)$ is the activation function of layer $l = 0, 1, \ldots, L$, and $\boldsymbol{\theta}^{(l)}$ is a weight matrix containing the weights and biases that connect the layer l with its successive layer, see Fig. 1. The use of nonlinear activation

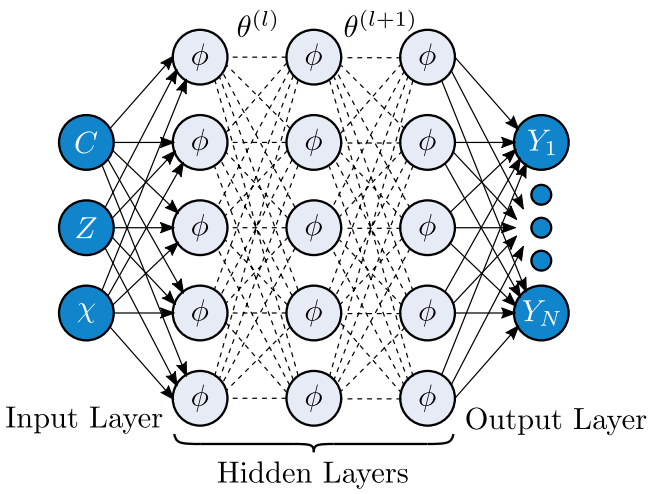


Fig. 1. Schematic Structure of the ANN used to approximate the mass fractions (Y_k) based on the input control variables (C, Z, and χ).

functions accounts for the non-linear and smooth interpolations of ANNs where – except for the final layer activation $\phi^{(L)}$ – common choices for regression tasks are the hyperbolic tangent (tanh) or Sigmoid linear unit (swish). For simplicity, all network parameters are from now on denoted with a single weight vector $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(L)})$. Also, ANN predictions of mass fractions are denoted by $\boldsymbol{Y}_{\boldsymbol{\theta}}(\boldsymbol{X})$, while $Y_{\boldsymbol{\theta},k}(\boldsymbol{X})$ denotes prediction of the mass fraction of species k.

Finding the optimal number of layers and neurons per layer, hence the optimal number of network parameters, can be quite cumbersome and is often based on trial-and-error. In general, increasing the number of network parameters increases the network's complexity, thus improving the capability of learning more complex reaction mechanisms. This number, however, also affects several stages of practically employing an ANN in a simulation, e.g., training time, memory utilization and inference time for real-time predictions. The training time scales with the number of network parameters, while memory utilization and inference time might be determined by the exact ANN coupling and used hardware. Furthermore, a greater number of network parameters makes the network more prone to overfitting the training data which, in case of low-quality data, significantly impacts the prediction accuracy in the interpolation regime.

Once the number of layers and neurons has been set, finding the optimal values of the network parameters is part of the training procedure that requires the selection of an appropriate loss function and optimization procedure. The standard loss function for regression tasks in machine learning is the well-known mean squared error (MSE) loss

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{Y}^{(i)} - \boldsymbol{Y}_{\boldsymbol{\theta}}(\boldsymbol{X}^{(i)})\right\|_2^2, \tag{2}$$

which is evaluated on examples $\left\{(\boldsymbol{X}^{(i)}, \boldsymbol{Y}^{(i)})\right\}_{i=1}^{N}$ in the training database. The optimization, i.e., finding the network parameters that minimize this loss function, is most commonly performed in an iterative and gradient-based process, according to the general update rule

$$\boldsymbol{\theta}_{n+1} \leftarrow \boldsymbol{\theta}_n - \alpha\frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}}, \tag{3}$$

where α denotes the learning rate, and n the optimization step that is usually repeated for a certain number of epochs. The gradients $\partial\mathcal{L}/\partial\boldsymbol{\theta}$ are calculated via the backpropagation algorithm as followed by common optimizers, such as Adam [28] or SGD [29].

3. Neural network approach

In this section, details on the ANN approach are provided, which includes the conservation of mass (CoM) network architecture, species loss weighting (SLW), as well as network and optimization settings, in that respective order.

3.1. Conservation of mass (CoM)

Conservation of mass is a fundamental physical law that plays a crucial role in closed chemical systems. Violation of this fundamental property would lead to an increase or decrease of total mass in the chemical system that accumulates over time, rendering simulation inaccurate and nonphysical. This can be critical for local species mass fractions, which are used in the energy equation to retrieve the local temperature. Violation of mass may propagate into temperature inaccuracies which thus render chemical reaction rates inaccurate. It is therefore of great interest to ensure in simulations that the mass of the system always remains constant, or in terms of mass fractions

$$\sum_k Y_k = 1, \tag{4}$$

which holds true for a complete reaction mechanism. Note that employing multiple ANNs, e.g., one for each species mass fraction, does not trivially fulfill CoM. Correction are thus often applied as a post-processing step which could increase the computational complexity and efforts needed for ANN training and predictions [16].

Using a single ANN, on the other hand, enables embedding CoM within the network structure. This can be achieved by taking the softmax activation function as final layer activation, which is often applied to approximate probability density functions and ensures that predictions sum up to one [30]. The softmax activation for the neuron k in the final layer L, which thus represents the approximated mass fraction $Y_{\boldsymbol{\theta},k}$ of species k, is given by

$$\phi^{(L)}(\boldsymbol{z})_k = \frac{e^{z_k}}{\sum_i e^{z_i}}, \tag{5}$$

where $\boldsymbol{z}$ is the input to the final layer, cf. Eq. (1). The normalization factor in the denominator of Eq. (5) ensures that any prediction of mass fractions fulfills

$$\sum_k Y_{\boldsymbol{\theta},k} = 1, \tag{6}$$

as stated by Eq. (4), and furthermore that individual species mass fractions lie in the physically meaningful range $Y_{\boldsymbol{\theta},k} \in (0, 1)$. An alternative and often used choice would be the Sigmoid activation function as final layer activation which indeed also ensures that individual predictions lie in the range $Y_{\boldsymbol{\theta},k} \in (0, 1)$, but they do not sum to one for multiple species. Throughout this work, both Sigmoid and softmax activations will be tested and referred to as the standard and CoM approach, respectively.

Algorithm 1 Pseudo code for ANN optimization using Species Loss Weighting (SLW)

Require: Y_k ▷ True data labels
Require: $\boldsymbol{\theta}, \alpha$ ▷ Initialize ANN & learning rate
$\omega_k \leftarrow SLW(Y_k)$ ▷ Set optimal species loss weights
while training **do**
  for each species k **do**
    $\mathcal{L}_k \leftarrow MSE(Y_k, Y_{\boldsymbol{\theta},k})$ ▷ Individual species loss
  end for
  $\mathcal{L} \leftarrow \sum_k \omega_k \mathcal{L}_k$ ▷ Final loss value
  $\boldsymbol{\theta}_{n+1} \leftarrow \boldsymbol{\theta}_n - \alpha \cdot \frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}}$ ▷ Gradient update
end while

3.2. Species loss weighting (SLW)

From a multi-variate regression perspective, learning multiple targets with heterogeneous scales poses challenges for the gradient-based optimization of a single ANN: by assuming equal relative prediction

errors across different targets, the optimization and respective gradients are potentially dominated by targets with larger values, as the MSE is bound to absolute prediction errors [31]. This is clear when the multivariate loss function in Eq. (2) is restated as the sum of individual species contributions

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{i=1}^{N}\sum_{k}\left(Y_k^{(i)} - Y_{\boldsymbol{\theta},k}(\boldsymbol{X}^{(i)})\right)^2 = \sum_{k}\mathcal{L}_k(\boldsymbol{\theta}), \tag{7}$$

where $\mathcal{L}_k$ denotes the MSE loss evaluated for species k. A detailed reaction mechanism with multiple species and heterogeneous target scales thus potentially results in minor species being learned less accurately than others, as the respective gradients $\partial\mathcal{L}_k/\partial\boldsymbol{\theta}$ in the general update rule (3) are also smaller compared to major species. This has been studied in [31], and suggests weighting the individual species components in the final loss function according to

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{k}\omega_k\mathcal{L}_k(\boldsymbol{\theta}), \tag{8}$$

where ω_k is a scalar loss weight for species k, and the use of this species weighted loss function is now referred to as the SLW approach. According to Eq. (8), individual gradients in the update rule (3) are then also scaled by the SLW approach. The loss weights ω_k intuitively adjust the importance of individual species in the optimization, which thus can be used to amplify gradients of underrepresented targets such as minor species mass fractions. Yet, the optimal choice of loss weights remains undetermined and is potentially influenced by the particular data distribution used in the training.

To keep the SLW approach as simple as possible, three different strategies are considered that individually take into account the distribution of target scales. These schemes are referred to as (i) MEAN, (ii) RANGE and (iii) VAR, representing which information is used for adjusting the loss weights. According to that, scheme (i) specifies the reciprocal of mean mass fractions $\omega_k := \overline{Y}_k^{-1}$ as loss weights, while (ii) uses its ranges $\omega_k := (\max\{Y_k\} - \min\{Y_k\})^{-1}$. To also incorporate higher moments of the underlying data distributions, scheme (iii) uses the variance $\mathrm{Var}(Y_k) = \sum_{i=1}^{N}(Y_k^{(i)} - \bar{Y}_k)^2/N$ to set the loss weights $\omega_k := \mathrm{Var}(Y_k)^{-1}$. The loss weights are determined on the species mass fractions in the training databases, and are subsequently kept fixed during the entire optimization process. Note that for inert species or species with vanishing variance in the mass fraction, loss weights and further their respective gradients in the optimization could explode when using the SLW approach. As a remedy for those cases, corrections could be made, e.g., by setting a maximum value for loss weights or adding a small number to the range or variance in scheme (ii) and (iii). To demonstrate the basic usage of the SLW approach, Algorithm 1 shows a pseudo code example.

Note that the SLW approach seems similar to performing target scaling in a data preprocessing step, which indeed it is but the SLW approach effectively applies the scaling within the ANN optimization procedure and thus works with unscaled targets. This enables the use of the softmax activation in the final output layer which would not be possible by considering target scaling. Furthermore, physical soft constraints, such as used by physics-informed neural networks [32,33], could be additionally applied since physical quantities are unmodified.

3.3. Data, network and optimizer settings

Four reaction mechanisms are generated with the flamelet model and serve as test case for the ANN application. These include detailed and skeletal (or reduced) mechanisms for H_2, $C_{12}H_{26}$, OME_{34}, and C_7H_{16}, which are studied as part of atmospheric burners (H_2) and compression-ignition engines ($C_{12}H_{26}$, OME_{34}, C_7H_{16}) configurations. Details are given in Table 1 where the list of involved species can be extracted from Fig. 4. The skeletal mechanisms include just a reduced number of species compared to the actual mechanism, where this is a common approach to reduce computational expenses. In order to compensate the disregard of mass in reducing the number of species, a single species is introduced that represents the omitted species and their combined mass fractions to still ensure CoM. This additional mass fraction is used for the $C_{12}H_{26}$, OME_{34}, and C_7H_{16} databases and subsequently also modeled by the ANN with the CoM architecture.

Table 1
Reaction mechanism, combustion settings and database numbers. The chamber temperature T_c and chamber pressure P_c define the combustion settings that were used in the atmospheric burner (H_2) and compression-ignition engines ($C_{12}H_{26}$, OME_{34}, C_7H_{16}) configurations. The last column represents the number of data points that is used for ANN training and testing.

Mechanism	Ref.	Species (reduced[a])	Reactions	T_c [K]	P_c [bar]	Data points
H_2	[34]	9	21	1045	1.0	6.6 mio.
$C_{12}H_{26}$	[35]	54 (15)	269	900	59.6	4.4 mio.
OME_{34}	[36]	322 (16)	1611	900	59.6	5.3 mio.
C_7H_{16}	[37]	248 (17)	1428	900	59.6	3.4 mio.

[a] Modeled in this work.

Table 2
Settings that where found to work well and used for the experiments in this work. The top part of hyperparameters are associated with the ANN settings, the lower part with the optimization procedure.

Setting	Value
Layers & Neurons[a]	2 × 20, 4 × 50
Activation (hidden)	tanh
Activation (output)	sigmoid (standard), softmax (CoM)
Initialization	Glorot
Optimizer	Adam
Learning Rate	0.001
Epochs	50
Batch Size	1024

[a] Different settings are used in Section 4.3.

In practice, the ANN would be trained on the full dataset, i.e., on all available data, to learn the reaction mechanisms as accurate as possible. Since this work only performs a prior assessment of the ANN performance, a standard splitting scheme is used to evaluate the ANN performance on a dedicated test set. For this, 10% of the data is randomly sampled from the full dataset for each reaction mechanism listed in Table 1, to provide the test set for the final performance evaluation. The remaining 90% represents the training data used for optimizing the ANN where in Section 4.1 this training data set is further used in a 5-fold cross validation scheme to determine the optimal choice of loss weights.

Results from previous experiments, showed that the settings listed in Table 2 work generally well on the presented tasks, where tested settings included the variation of layers (1,2,3,4,5), neurons per layer (20,50,70), activation (tanh, sigmoid, swish), initialization (Glorot, He), optimizers (Adam, SGD), learning rates (0.01, 0.001, 0.0001), epochs (50,100,150) and batch sizes (64,1024,8192). Since qualitative conclusions were the same across several tests, the default settings in Table 2 are used in Sections 4.1 and 4.2, while different layer settings are used in Section 4.3.

4. Results

This section is structured as follows: First, the different species loss weighting schemes (discussed in Section 3.2) are evaluated to set the optimal and preferred choice of loss weights for the subsequent experiments. Next, both modifications SLW and CoM (see Section 3) are compared to the standard use of a single ANN by measuring the performance on individual species mass fractions. In the final part, the trade-off between accuracy and ANN complexity is evaluated by comparing the proposed approach with the common application of using multiple ANNs.

To validate the ANN performance, the Pearson product-moment correlation coefficient (PPMCC) is selected which is often used for

Table 3
Mean (and standard deviation) of the overall ANN performances $\overline{\rho}$ across five uniquely trained ANN instances. The overall ANN performance is determined by the average of all species performances ρ_k. **Bold:** Best performance for each reaction mechanism.

	Network Architecture, 2 × 20			
SLW	H_2	$C_{12}H_{26}$	OME_{34}	C_7H_{16}
Standard ANN	0.648 (0.485)	0.666 (0.436)	0.679 (0.455)	0.599 (0.448)
MEAN	0.987 (0.027)	0.975 (0.029)	0.982 (0.021)	0.908 (0.232)
RANGE	0.979 (0.043)	0.952 (0.065)	0.962 (0.061)	0.829 (0.293)
VAR	**0.999 (0.001)**	**0.985 (0.011)**	**0.990 (0.008)**	**0.968 (0.033)**
	Network Architecture, 4 × 50			
SLW	H_2	$C_{12}H_{26}$	OME_{34}	C_7H_{16}
Standard ANN	0.826 (0.355)	0.867 (0.307)	0.880 (0.299)	0.764 (0.399)
MEAN	0.999 (0.001)	0.993 (0.010)	0.997 (0.005)	0.873 (0.321)
RANGE	0.997 (0.007)	0.988 (0.019)	0.993 (0.012)	0.845 (0.330)
VAR	**1.000 (0.001)**	**0.997 (0.003)**	**0.998 (0.002)**	**0.987 (0.018)**

quantifying the a priori prediction accuracy in combustion studies [14, 38]. The PPMCC depicts the linear correlation between two sets of data and, in the context of this work, is calculated for species k according to

$$\rho_k = \frac{\sum_{i=1}^{N}\left(Y_k^{(i)} - \overline{Y}_k\right)\left(Y_{k,\theta}^{(i)} - \overline{Y}_{k,\theta}\right)}{\sqrt{\sum_{i=1}^{N}\left(Y_k^{(i)} - \overline{Y}_k\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_{k,\theta}^{(i)} - \overline{Y}_{k,\theta}\right)^2}}, \tag{9}$$

where $\rho_k \in [-1, 1]$ and $\rho_k = 1$ indicates a perfect fit between true and predicted mass fractions.

4.1. Optimal species loss weighting

In the following experiments, different schemes for the SLW approach are evaluated and compared to a standard ANN that does not use the SLW approach. In order to produce comparable results, all ANNs use the standard sigmoid activation as final layer activation, i.e. predicted mass fractions are bound with $Y_{\theta,k} \in (0, 1)$ but the CoM architecture is not applied. Two ANN settings are used, namely a small network architecture with 2 × 20 and a larger one with 4 × 50. For each setting 5-fold cross validation is applied, i.e. five repetitions with different seeds for network initialization and data splits are performed.

The results are listed in Table 3, which shows the mean (and standard deviation) of the overall performance $\overline{\rho}$ that is determined by the average of all species performances ρ_k. As evident from these numbers, the standard ANN only achieves a moderate performance on each reaction mechanism. The standard ANN performance is $\overline{\rho} < 0.7$ for the small network architecture and $\overline{\rho} < 0.88$ for the larger one, where the lowest performance for both cases is measured on the C_7H_{16} reaction mechanism. Further investigations have shown that these numbers are mainly influenced by minor species performances (will be shown in the next experiments). Note that while testing various ANN and optimizer settings, no significant improvement could be achieved with the standard ANN approach. The performance issues related to minor species appeared to be unaffected by settings other than those listed in Table 2, why the overall performance of the standard ANN could not be further improved considerably. The use of the SLW approach, on the other hand, greatly improves the overall performance on each reaction mechanisms as it is evident from Table 3. Among all SLW schemes, the VAR scheme seems to be the best choice for the loss weights, as it consistently outperforms the other schemes and achieves an overall performance of $\overline{\rho} > 0.96$ even with the small network architecture. The superiority of the VAR scheme is believed to be explained by the fact that this scheme effectively normalizes individual species to unit variance. This appears beneficial for training with an MSE criterion in the followed multivariate regression setting. It can be expected that for other optimization criteria, such as mean absolute error, alternative weight scaling approaches may be more effective.

To visually demonstrate the effectiveness of the SLW approach, individual species predictions on the test sets are provided by using the 4 × 50 network architecture and all available data from the training sets. Fig. 2 and 3 show the correlation plots for a selection of major and minor species, namely H_2, HO_2, and H_2O_2 from the H_2 database, and $C_{12}H_{26}$, O, and H from the $C_{12}H_{26}$ database. The plots show the true versus the predicted mass fractions, where the diagonal line indicates a perfect fit with $\rho_k = 1$. The respective accuracies are given in the upper left corner of each plot. Evident from the figures is that all approaches perform well on the major species (H_2 and $C_{12}H_{26}$) as indicated by a nearly perfect fit with $\rho_k \approx 1$. However, for the minor species (HO_2, H_2O_2, O, H) the performance of the standard ANN approach is poor and seems to decline with the species depletion. The MEAN and RANGE schemes perform considerably better on the minor species while the VAR scheme achieves the best results, in particular considering the outstanding performance on HO_2 and H_2O_2 in the H_2 database in Fig. 2. Based on these results the VAR method will be used for all remaining experiments.

4.2. ANN performance on individual species

As a continuation of the preceding experiment, the prediction accuracy of all individual species in the reaction mechanisms is evaluated. For comparison, the following three ANN strategies are applied: a standard ANN is employed that serves as reference and is set up as discussed in the previous experiment. To now also study the effectiveness of the CoM architecture, the SLW approach with the VAR scheme is employed with two settings, once with the sigmoid activation (without CoM) and once with the softmax activation (with CoM) in the final output layer. The three approaches are named in the respective order as standard ANN, SLW, and SLW+CoM. For demonstration purposes, the 2 × 20 network architecture is used since the 4 × 50 architecture has already exhibited outstanding performance with the SLW approach in the preceding experiment — the difference between the SLW and SLW+CoM would be barely noticeable with the 4 × 50 network architecture. The remaining settings are taken as listed in Table 2. The training is again performed on the complete training dataset where predictions are made on the dedicated test set, respectively for each reaction mechanism.

The results are shown in Fig. 4 where each subplot represents a single reaction mechanism. The order of listed species is based on the performance of the standard ANN approach, decreasing from left to right. Across all reaction mechanisms, it is clearly evident that the performance of the standard ANN declines with the depletion of species, i.e. minor species with small mass fractions are predicted worse than major species. As already captured by Fig. 2 and 3, the standard ANN fails largely ($\rho_k < 0$) in predicting the mass fractions of certain minor species, in particular that of H and H_2O_2 in the H_2 database, and O, H, and OH in the others. The SLW approach, on the other hand, yields accurate results across all species and databases, demonstrating its robustness for distinct reaction mechanisms. This was already captured by the overall performance in Table 3. From the figure it is now also evident that the additional use of the CoM architecture further improves the performance on every single species — again, the difference is less pronounced with the 4 × 50 network architecture (not shown) but the same conclusion can be derived. In view of this, the additional constraint of ensuring that the sum of all species mass fractions equals one not only guarantees CoM when utilizing the ANN in simulations but also appears to have a generally favorable impact on the performance of individual species. This effectively demonstrates the benefits of incorporating simple physical constraints into ANNs — notably, with the CoM architecture, these benefits are achieved without the need for any additional computations.

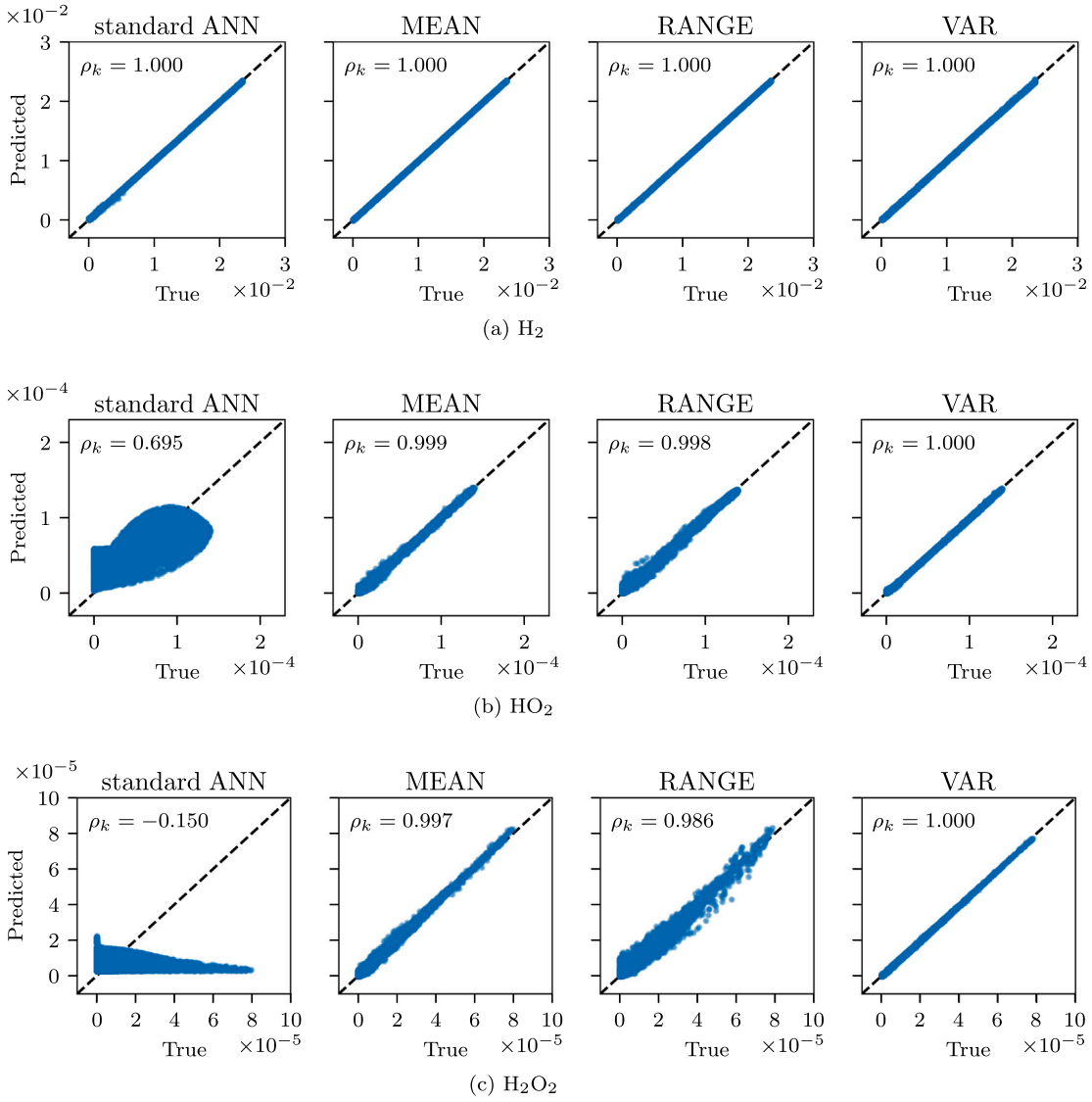


Fig. 2. True vs. predicted mass fractions for species (a) H_2, (b) HO_2, and (c) H_2O_2 from the H_2 test data using the 4×50 network architecture. A perfect fit ($\rho_k = 1$) is indicated by the dashed diagonal line.

4.3. Computational efficiency & scalability

As previous results showed that the proposed ANN approach is capable of accurately learning all species mass fraction in a complete reaction mechanism, the approach is put into a final test to evaluate its computational efficiency and scalability. For this, the trade-off between ANN accuracy and network complexity is assessed. The proposed approach (SLW+CoM) is compared with the method of utilizing multiple ANNs (one for each species; MMLP [20,21]), which is the only solution known from the literature that has also demonstrated accurate predictions for minor species, see Section 1. In order to provide a fair evaluation of their computational efficiency, the ANN performance is evaluated for different numbers of layers L, by focusing on the performance of the least accurately learned species. This performance usually determines whether an ANN approach can be successfully employed in a chemical combustion simulation (discussed in Section 2.2). The used setup thus evaluates the computational efficiency of the ANN approach for several stages of the ANN utilization, such as training time and memory utilization.

Tests are performed on the H_2 reaction mechanism with 20 neurons per layer and on the more complex C_7H_{16} reaction mechanism with 50 neurons per layer. The number of layers, on the other hand, is varied from one to five layers to adjust the number of network parameters and thus the ANN complexity. The MMLP approach is used for the complete reaction mechanisms, i.e., in total 9 ANNs for the H_2 system and 16 ANNs for the C_6H_{16} system; one ANN per species. Furthermore, the sigmoid activation[1] is applied in the final output layer of each ANN in the MMLP approach, and MinMax target scaling is used to optimize all ANNs with an equal learning rate of 0.001.

The results for the two systems are shown in Fig. 5(a) and 5(b). In the two figures, the solid line (left y-axis) shows the maximum prediction error which is based on the least accurately learned species in the complete reaction mechanism and determined by $1 - \min\{\rho_k\}$. A lower prediction error implies higher chances of success in employing the ANN in a simulation. The dashed line (right y-axis) captures the number of network parameters θ, i.e. the sum of all ANN weights and biases used by the respective approach. From the figures it is evident that the SLW+CoM approach is computationally more efficient compared to the

[1] Conservation of mass is disregarded in this experiment, but would potentially also influence the ANN's performance in a simulation.

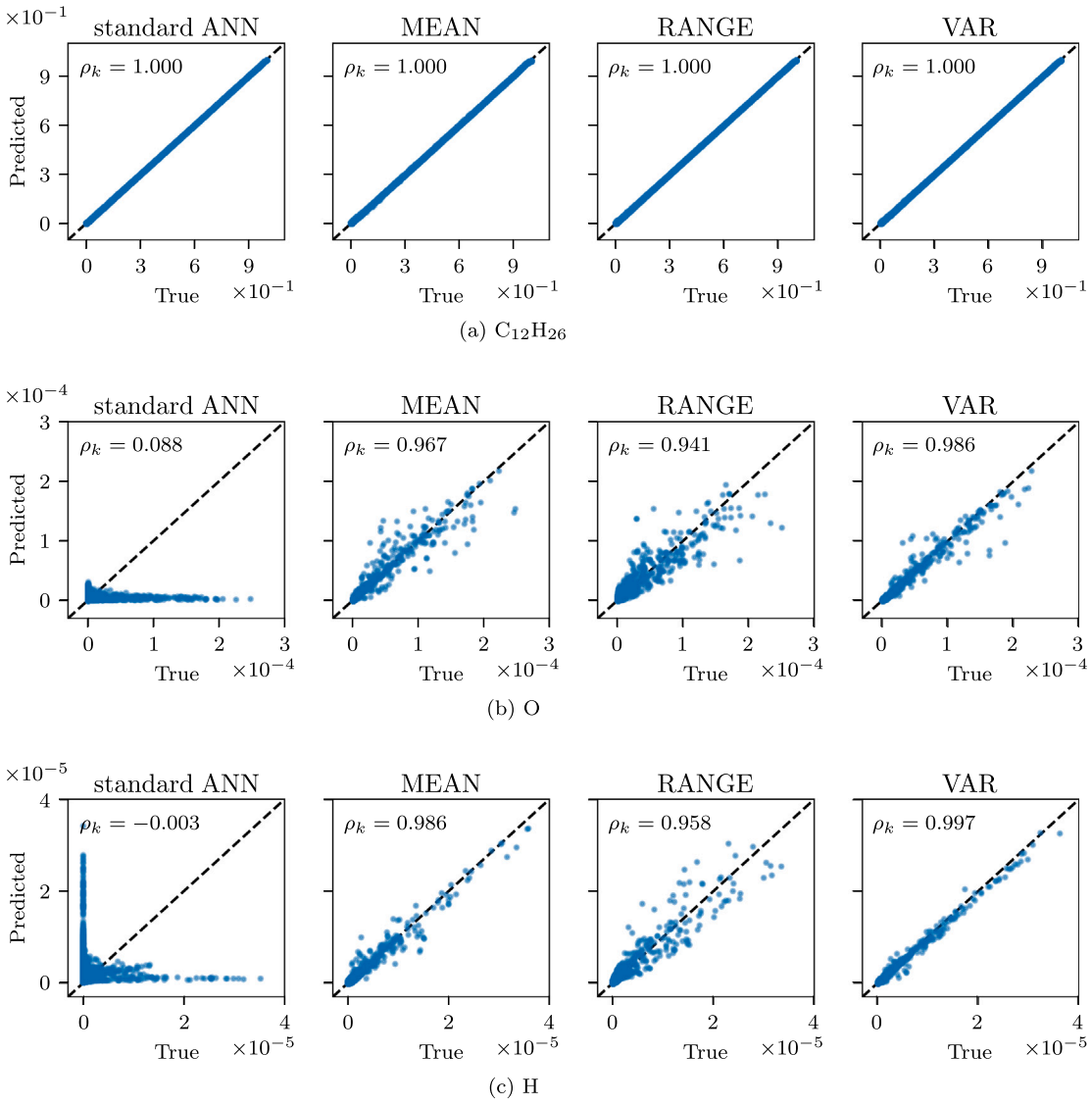


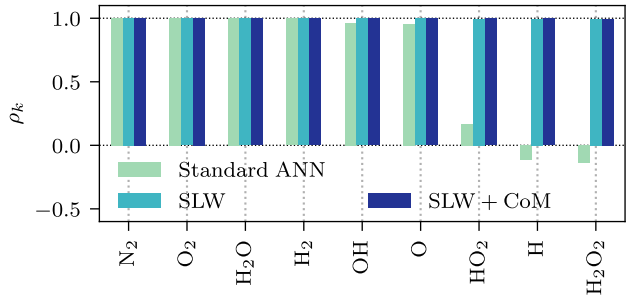
Fig. 3. True vs. predicted mass fractions for species (a) $C_{12}H_{26}$, (b) O, and (c) H from the $C_{12}H_{26}$ test data using the 4×50 network architecture. A perfect fit ($\rho_k = 1$) is indicated by the dashed diagonal line.

MMLP approach, as fewer network parameters are needed to achieve a low prediction error. For instance, both approaches achieve a low prediction error with five layers. The SLW+CoM approach, however, only needs a fraction of network parameters in comparison to the MMLP approach.[2] Furthermore, comparing Fig. 5(a) and 5(b), the difference of used network parameters seems to become more distinct as the reaction mechanism involves more species. This is due to the fact that the MMLP approach scales linearly with the number of involved species while the SLW+CoM does not.

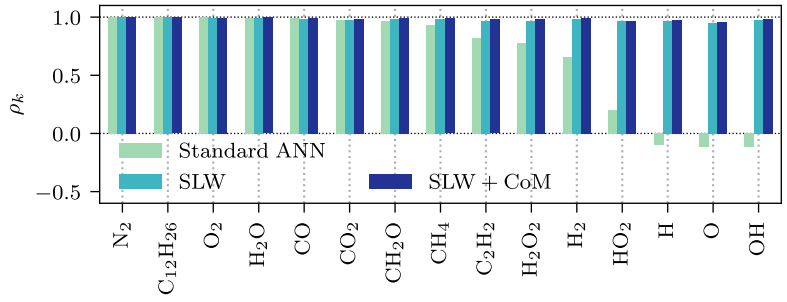
5. Conclusion

This paper presented a novel ANN approach designed for modeling chemical kinetics and species mass fractions as used by tabulation methods in combustion studies. The proposed approach leverages a network architecture that conserves mass, and utilizes a particular weighting of species depletion. Both simple modifications have proven effective in accurately learning and predicting multiple mass fractions with a single ANN. The efficacy of the approach was validated through experiments involving four distinct reaction mechanisms (H_2, C_7H_{16}, $C_{12}H_{26}$, OME_{34}), with a comparative analysis against standard ANN approaches. Results showed that incorporating mass conservation led to an enhancement in overall performance, and the introduced species loss weighting significantly improved the prediction accuracy of individual species mass fractions. Notably, even with multiple minor species involved in the reaction mechanism, the ANN approach demonstrated robust performance, underscoring its capability to handle widely varying target scales. Additionally, an assessment of computational efficiency and scalability to complex reaction mechanisms was conducted, revealing the superior performance of the proposed approach compared to standard ANN approaches.

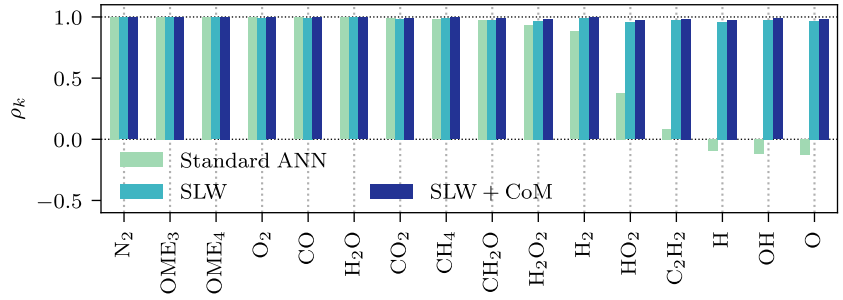
[2] For the used setup and hardware, training a single 2×20 ANN instance took about 40 min; multiplied by the number of species for the MMLP approach.



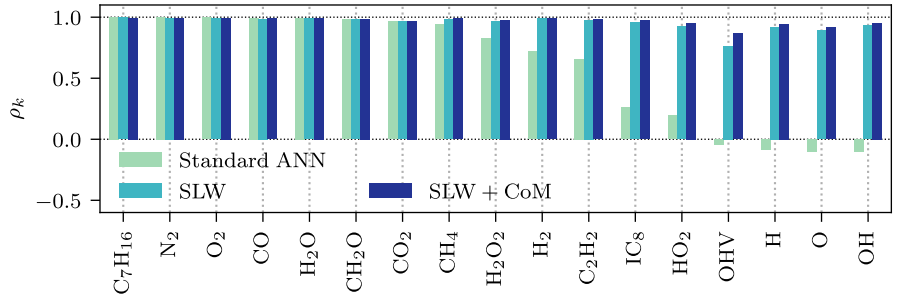
(a) H_2 database



(b) $C_{12}H_{26}$ database

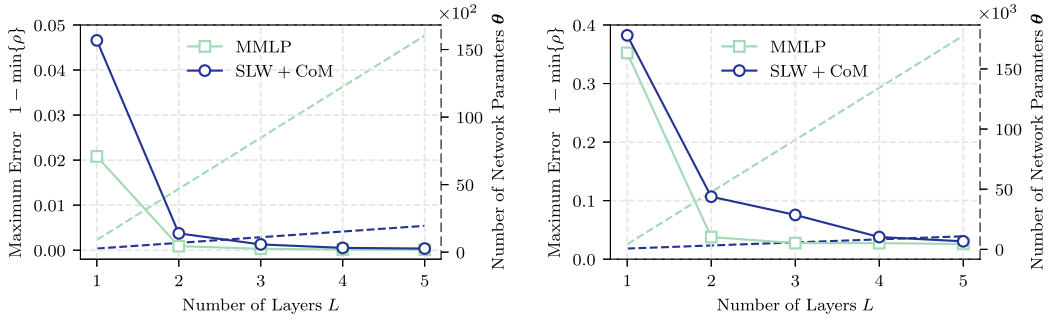


(c) OME_{34} database



(d) C_7H_{16} database

Fig. 4. Performance on individual species evaluated on the test sets for the (a) H_2, (b) $C_{12}H_{26}$, (c) OME_{34}, and (d) C_7H_{16} database using the 2×20 network architecture.



(a) H_2 database (b) C_7H_{16} database

Fig. 5. Computational efficiency of the MMLP and SLW+CoM approach, evaluated for different numbers of layers L on the (a) H_2 and (c) C_7H_{16} database. The solid line (left y-axis) shows the maximum prediction error across all species in the reaction mechanism, which is given by $1 - \min\{\rho\}$ and thus selected by the lowest species performance ρ_k. The dashed line (right y-axis) shows the number of network parameters θ applied by the respective approach.

CRediT authorship contribution statement

Franz M. Rohrhofer: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Stefan Posch:** Conceptualization, Formal analysis, Resources, Supervision, Writing – review & editing . **Clemens Gößnitzer:** Supervision, Writing – review & editing . **José M. García-Oliver:** Data curation, Resources, Writing – review & editing . **Bernhard C. Geiger:** Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Franz M. Rohrhofer reports financial support was provided by Austrian Research Promotion Agency.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported by the Austrian COMET — Competence Centers for Excellent Technologies — Programme of the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology, the Austrian Federal Ministry for Digital and Economic Affairs, and the States of Styria, Upper Austria, Tyrol, and Vienna for the COMET Centers Know-Center and LEC EvoLET, respectively. The COMET Programme is managed by the Austrian Research Promotion Agency (FFG).

References

[1] Zhou L, Song Y, Ji W, Wei H. Machine learning for combustion. Energy AI 2022;7:100128. http://dx.doi.org/10.1016/j.egyai.2021.100128.

[2] An J, He G, Luo K, Qin F, Liu B. Artificial neural network based chemical mechanisms for computationally efficient modeling of hydrogen/carbon monoxide/kerosene combustion. Int J Hydrogen Energy 2020;45(53):29594–605. http://dx.doi.org/10.1016/j.ijhydene.2020.08.081.

[3] Flemming F, Sadiki A, Janicka J. LES using artificial neural networks for chemistry representation. Prog Comput Fluid Dyn, Int J 2005;5(7):375–85. http://dx.doi.org/10.1504/PCFD.2005.007424.

[4] Amini Y, Fattahi M, Khorasheh F, Sahebdelfar S. Neural network modeling the effect of oxygenate additives on the performance of Pt–Sn/γ-Al 2 O 3 catalyst in propane dehydrogenation. Appl Petrochem Res 2013;3:47–54. http://dx.doi.org/10.1007/s13203-013-0028-8.

[5] An J, Qin F, Zhang J, Ren Z. Explore artificial neural networks for solving complex hydrocarbon chemistry in turbulent reactive flows. Fund Res 2022;2(4):595–603. http://dx.doi.org/10.1016/j.fmre.2021.08.007.

[6] Fattahi M, Kazemeini M, Khorasheh F, Rashidi A. Kinetic modeling of oxidative dehydrogenation of propane (ODHP) over a vanadium–graphene catalyst: Application of the DOE and ANN methodologies. J Ind Eng Chem 2014;20(4):2236–47. http://dx.doi.org/10.1016/j.jiec.2013.09.056.

[7] Amato F, González-Hernández JL, Havel J. Artificial neural networks combined with experimental design: A "soft" approach for chemical kinetics. Talanta 2012;93:72–8. http://dx.doi.org/10.1016/j.talanta.2012.01.044.

[8] Van Oijen J, De Goey L. Modelling of premixed laminar flames using flamelet-generated manifolds. Combust Sci Technol 2000;161(1):113–37. http://dx.doi.org/10.1080/00102200008935814.

[9] Bekdemir C, Somers L, De Goey L. Modeling diesel engine combustion using pressure dependent flamelet generated manifolds. Proc Combust Inst 2011;33(2):2887–94. http://dx.doi.org/10.1016/j.proci.2010.07.091.

[10] Chatzopoulos A, Rigopoulos S. 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 2013;34(1):1465–73. http://dx.doi.org/10.1016/j.proci.2012.06.057.

[11] Elbahloul S, Rigopoulos S. Rate-Controlled Constrained Equilibrium (RCCE) simulations of turbulent partially premixed flames (Sandia D/E/F) and comparison with detailed chemistry. Combust Flame 2015;162(5):2256–71. http://dx.doi.org/10.1016/j.combustflame.2015.01.023.

[12] Kohonen T. Essentials of the self-organizing map. Neural Netw 2013;37:52–65. http://dx.doi.org/10.1016/j.neunet.2012.09.018.

[13] Blasco JA, Fueyo N, Dopazo C, Chen J. A self-organizing-map approach to chemistry representation in combustion applications. Combust Theory Model 2000;4(1):61. http://dx.doi.org/10.1088/1364-7830/4/1/304.

[14] Zhang Y, Xu S, Zhong S, Bai X-S, Wang H, Yao M. Large eddy simulation of spray combustion using flamelet generated manifolds combined with artificial neural networks. Energy AI 2020;2:100021. http://dx.doi.org/10.1016/j.egyai.2020.100021.

[15] Honzawa T, Kai R, Hori K, Seino M, Nishiie T, Kurose R. Experimental and numerical study of water sprayed turbulent combustion: Proposal of a neural network modeling for five-dimensional flamelet approach. Energy AI 2021;5:100076. http://dx.doi.org/10.1016/j.egyai.2021.100076.

[16] Readshaw T, Jones WP, Rigopoulos S. On the incorporation of conservation laws in machine learning tabulation of kinetics for reacting flow simulation. Phys Fluids 2023;35(4). http://dx.doi.org/10.1063/5.0143894.

[17] Peng WY, Pinkowski NH. Efficient and accurate time-integration of combustion chemical kinetics using artificial neural networks. 2017.

[18] Blasco J, Fueyo N, Dopazo C, Ballester J. Modelling the temporal evolution of a reduced combustion chemical system with an artificial neural network. Combust Flame 1998;113(1–2):38–52. http://dx.doi.org/10.1016/S0010-2180(97)00211-3.

[19] Kempf A, Flemming F, Janicka J. Investigation of lengthscales, scalar dissipation, and flame orientation in a piloted diffusion flame by LES. Proc Combust Inst 2005;30(1):557–65. http://dx.doi.org/10.1016/j.proci.2004.08.182.

[20] Ding T, Readshaw T, Rigopoulos S, Jones W. Machine learning tabulation of thermochemistry in turbulent combustion: An approach based on hybrid flamelet/random data and multiple multilayer perceptrons. Combust Flame 2021;231:111493. http://dx.doi.org/10.1016/j.combustflame.2021.111493.

[21] Ding T, Rigopoulos S, Jones W. Machine learning tabulation of thermochemistry of fuel blends. Appl Energy Combust Sci 2022;12:100086. http://dx.doi.org/10.1016/j.jaecs.2022.100086.

[22] Peters N. Laminar diffusion flamelet models in non-premixed turbulent combustion. Prog Energy Combust Sci 1984;10(3):319–39. http://dx.doi.org/10.1016/0360-1285(84)90114-X.

[23] Naud B, Novella R, Pastor JM, Winklinger JF. RANS modelling of a lifted H2/N2 flame using an unsteady flamelet progress variable approach with presumed PDF. Combust Flame 2015;162(4):893–906. http://dx.doi.org/10.1016/j.combustflame.2014.09.014, URL http://linkinghub.elsevier.com/retrieve/pii/S0010218014002831.

[24] Pandal A, Garcia-Oliver JM, Novella R, Pastor JM. A computational analysis of local flow for reacting Diesel sprays by means of an Eulerian CFD model. Int J Multiph Flow 2018;99:257–72. http://dx.doi.org/10.1016/j.ijmultiphaseflow.2017.10.010.

[25] Desantes JM, Garcia-Oliver JM, Novella R, Pachano L. A numerical study of the effect of nozzle diameter on diesel combustion ignition and flame stabilization. Int J Engine Res 2020;21(1):101–21. http://dx.doi.org/10.1177/1468087419864203, URL http://journals.sagepub.com/doi/10.1177/1468087419864203.

[26] Ashraf WM, Uddin GM, Arafat SM, Krzywanski J, Xiaonan W. Strategic-level performance enhancement of a 660 MWe supercritical power plant and emissions reduction by AI approach. Energy Convers Manage 2021;250:114913. http://dx.doi.org/10.1016/j.enconman.2021.114913.

[27] Krzywanski J, Czakiert T, Nowak W, Shimizu T, Zylka A, Idziak K, et al. Gaseous emissions from advanced CLC and oxyfuel fluidized bed combustion of coal and biomass in a complex geometry facility: A comprehensive model. Energy 2022;251:123896. http://dx.doi.org/10.1016/j.energy.2022.123896.

[28] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014, arXiv preprint arXiv:1412.6980.

[29] Bottou L, et al. Stochastic gradient learning in neural networks. Proc Neuro-Nımes 1991;91(8):12.

[30] Goodfellow I, Bengio Y, Courville A. Softmax units for multinoulli output distributions. Deep Learning. MIT Press; 2018.

[31] Rohrhofer FM, Posch S, Gößnitzer C, García-Oliver JM, Geiger BC. Bringing chemistry to scale: Loss weight adjustment for multivariate regression in deep learning of thermochemical processes. 2023, arXiv preprint arXiv:2308.01954.

[32] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J Comput Phys 2019;378:686–707. http://dx.doi.org/10.1016/j.jcp.2018.10.045.

[33] Rohrhofer FM, Posch S, Gößnitzer C, Geiger BC. On the apparent Pareto front of physics-informed neural networks. IEEE Access 2023. http://dx.doi.org/10.1109/ACCESS.2023.3302892.

[34] Saxena P, Williams FA. Testing a small detailed chemical-kinetic mechanism for the combustion of hydrogen and carbon monoxide. Combust Flame 2006;145(1–2):316–23. http://dx.doi.org/10.1016/j.combustflame.2005.10.004.

[35] Yao T, Pei Y, Zhong B-J, Som S, Lu T, Luo KH. A compact skeletal mechanism for n-dodecane with optimized semi-global low-temperature chemistry for diesel engine simulations. Fuel 2017;191:339–49. http://dx.doi.org/10.1016/j.fuel.2016.11.083.

[36] Cai L, Jacobs S, Langer R, vom Lehn F, Heufer KA, Pitsch H. Auto-ignition of oxymethylene ethers (OMEn, n=2–4) as promising synthetic e-fuels from renewable electricity: shock tube experiments and automatic mechanism generation. Fuel 2020;264:116711. http://dx.doi.org/10.1016/j.fuel.2019.116711.

[37] Desantes JM, Garcia-Oliver JM, Novella R, Pastor J, Pintor DL, Shang W. Application of an optimized mechanism of primary reference fuel to single hole sprays. Tech. rep., SAE Technical Paper; 2023.

[38] Li K, Rahnama P, Novella R, Somers B. Combining flamelet-generated manifold and machine learning models in simulation of a non-premixed diffusion flame. Energy AI 2023;14:100266. http://dx.doi.org/10.1016/j.egyai.2023.100266.