Combustion and Flame 223 (2021) 142–152

Contents lists available at ScienceDirect

Combustion and Flame

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

A data-based hybrid model for complex fuel chemistry acceleration at high temperatures

Sultan Alqahtani[a,b,*], Tarek Echekki[a]

[a] Department of Mechanical and Aerospace Engineering, North Carolina State University, 1840 Entrepreneur Drive, Campus Box 7910, Raleigh, NC 27695, United States

[b] Department of Mechanical Engineering, King Khalid University, Abha, Saudi Arabia

ARTICLE INFO

Article history:
Received 7 April 2020
Revised 19 September 2020
Accepted 20 September 2020
Available online 14 October 2020

Keywords:
Data-based hybrid chemistry model
PCA
ANN

ABSTRACT

During their high-temperature oxidation, complex hydrocarbons and their early fragments are short-lived and figure prominently only during the pyrolysis stage. However, they are quickly replaced by smaller hydrocarbons at the onset of the oxidation stage, resulting in simpler chemistry requirements past pyrolysis. In this study, we develop a data-based hybrid chemistry approach to accelerate chemistry integration for complex fuels. The approach is based on tracking the evolution of chemistry through representative species for the pyrolysis and coupling their reactions with simpler foundational chemistry. The selection of these representative species is implemented using principal component analysis (PCA) based on simulation data. The description of chemistry for the representative species is implemented using an artificial neural network (ANN) model for their reaction rates followed by the description of their chemistry using a foundational chemistry model. The selection of the transition between these models is trained *a priori* using an ANN pattern recognition classifier. This data-based hybrid chemistry acceleration model is demonstrated for three fuels: n-dodecane, n-heptane and n-decane and investigated with two foundational chemistry, C_0–C_2 and C_0–C_4, models. The hybrid scheme results in computational saving, up to one order of magnitude for n-dodecane, two orders of magnitudes for n-heptane, and three orders of magnitudes for n-decane. The accuracy and saving in computational cost depend on the number of selected species and the size of the used foundational chemistry. The hybrid model coupled with the more detailed C_0–C_4 foundational performs, overall, better than the one coupled with the C_0–C_2 foundational chemistry.

1. Introduction

The integration of the chemistry of complex fuels is computationally challenging for simulations beyond 0D. Therefore, chemistry reduction or other strategies to accelerate chemistry integration have been an active area of research over the past 4 decades [1]. Strategies to accelerate chemistry include *in situ* adaptive tabulation (ISAT) [2] and regression (e.g. the piecewise reusable implementation for solution mapping or PRISM [3]), manifold-based methods, such as intrinsic low-dimensional manifolds (ILDM) [4] and computational singular perturbation (CSP) [5], and adaptive chemistry methods, including dynamic approaches (see for example Refs. [6–8]).

Additional strategies to accelerate chemistry can be adopted under certain mixture or combustion conditions. During high-temperature oxidation in complex hydrocarbon fuels, there are distinct stages associated with a fast pyrolysis followed by “slower” oxidation [9]. This behavior has enabled a few strategies to describe complex fuel chemistry, including the recent hybrid chemistry (HyChem) approach by Wang and co-workers [10–16]. HyChem relies on fuel/fragment measurements to develop a reduced chemistry description for fuels without the prior requirement for an existing detailed or reduced chemistry description for these fuels. In this approach, the global reactions for the fuel fragments are coupled with established foundational chemistry that models the oxidation of C_0–C_4 species.

Recently, we have proposed a similar approach to HyChem where the temporal profiles of measured species can be used directly to determine their chemical reaction rates [17,18]. In Ref. [18], we proposed that soon after the pyrolysis stage, the foundational chemistry can be further simplified to account primarily

* Corresponding author at: Department of Mechanical and Aerospace Engineering, North Carolina State University, 1840 Entrepreneur Drive, Campus Box 7910, Engineering Building III, Room 3252, Raleigh, NC 27695, United States.
E-mail addresses: ssalqaht@ncsu.edu (S. Alqahtani), techekk@ncsu.edu (T. Echekki).

https://doi.org/10.1016/j.combustflame.2020.09.022
0010-2180/

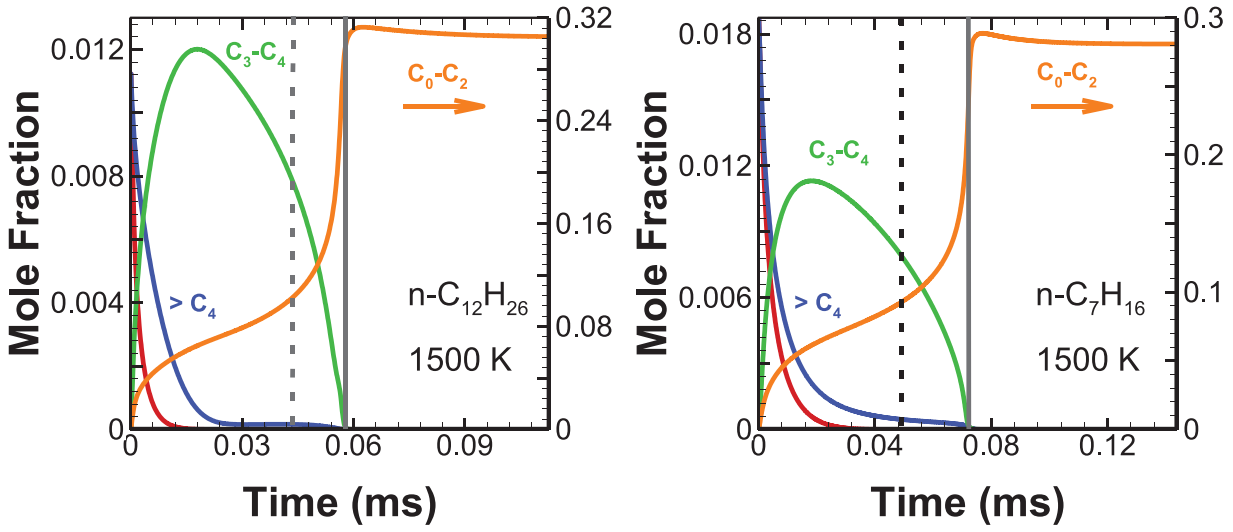

Fig. 1. Species temporal profiles from detailed homogenous chemistry simulations of n-dodecane and n-heptane/air stoichiometric mixtures at initial temperatures 1500 K and 1 atm pressure.

for C_0–C_2 species. The hierarchical development during complex fuel oxidation from complex to simpler hydrocarbons has been a key observation on which the HyChem approach and our approach have been based. In the section below, we will illustrate this observation for two fuels considered in this study.

The hybrid chemistry models proposed in [17,18] are inspired by the HyChem approach [10–17], with some key distinctions. Both models combine a reaction rate models for the fuel fragments with foundational chemistry for the remaining species. The HyChem approach [10–17] uses global reaction steps for the fragments whose rate parameters are obtained through an optimization process to capture key observables, such as ignition delay times. In contrast, Refs. [17,18] derive these reaction rates directly from temporal profiles of multi-scalar measurements using shallow artificial neural networks (ANNs). Moreover, the reaction rates for the fragments are modeled using ANNs in terms of the modeled fragments. The chemistry for a subset of these fragments may not be captured by a simpler foundational C_0–C_4 chemistry. The same set of strategies, of course, can be implemented for chemistry reduction/acceleration when an adequate (e.g. detailed) chemistry description is available. The use of such mechanisms may not be practical beyond 0D or 1D simulations of relevance to practical applications. Moreover, important questions remain as related to whether fragments need to be modeled within a hybrid chemistry framework or whether there are alternative species that can be used instead within this framework.

An additional reduction of the complexity of fuel oxidation and a potential impetus for chemistry acceleration is associated with the adoption of a reduced chemistry description under certain combustion conditions. As in HyChem and our hybrid chemistry approach, this can be achieved based on available data for these conditions. Indeed, the present study explores this and attempts to develop strategies for chemistry acceleration for the high-temperature oxidation of complex fuels within a data-based framework. Principal component analysis (PCA) a fundamental tool for dimensionality reduction is designed to identify representative species that can be used to track and integrate the evolution of the chemical system instead of the fuel fragments.

The objective of the present study is to develop and investigate such an approach. In the next section, Section 2, we provide a brief motivation for the proposed approach and discuss the implementation steps. In Section 3, results of the model's predictions are presented for 3 different fuels, n-dodecane, n-heptane and n-decane, using two foundational chemistry mechanisms and compared against predictions from detailed chemical mechanisms. In Section 4, the results are summarized, and future extensions of the present work are discussed.

2. Motivation and model implementation

In this section, we provide a motivation for the present approach followed by the model implementation.

2.1. Motivation for the approach

As discussed in our recent work [18], the hierarchical evolution in the complexity of the fuels representing the pyrolysis and oxidation stages provides a mechanism for simplifying the foundational chemistry within a hybrid chemistry framework. This evolution is illustrated in Fig. 1 for the high-temperature chemistry of stoichiometric mixtures in air of n-dodecane and n-heptane at an initial temperature of 1500 K and 1 atm. Hydrocarbons are grouped based on their carbon composition to distinguish groups, C_0–C_2 (orange lines), C_3–C_4 (green lines) and C_5 and higher ($> C_4$) (blue). The fuel, indicated in red lines, decays much faster to produce fragments. The vertical dashed lines delineate the transition in time to primarily C_0–C_4 chemistry, while the vertical solid lines delineate the transition in time to primarily C_0–C_2 chemistry. The figure clearly shows a hierarchical evolution in the complexity of the hydrocarbons as the oxidation proceeds in time. The strategy that we will adopt in this work exploits this evolution.

2.2. Model formulation

In our recent study [18], we proposed a formulation for chemistry acceleration for complex fuels at high-temperature based on a hybrid chemistry model that combines the chemistry of fuel fragments with a simpler foundational chemistry based on a C_0–C_2 mechanism. The fuel fragments' chemistry is modeled using a regression based on artificial neural networks (ANN) to relate the fragments' reaction rates to their concentrations starting with their measured temporal profiles. The reliance on temporal profiles is motivated by their potential availability from experimental measurements.

In the present approach, we propose a model to accelerate chemistry starting from detailed or skeletal mechanisms for complex fuels. Moreover, in contrast with the model in Ref. [18] where

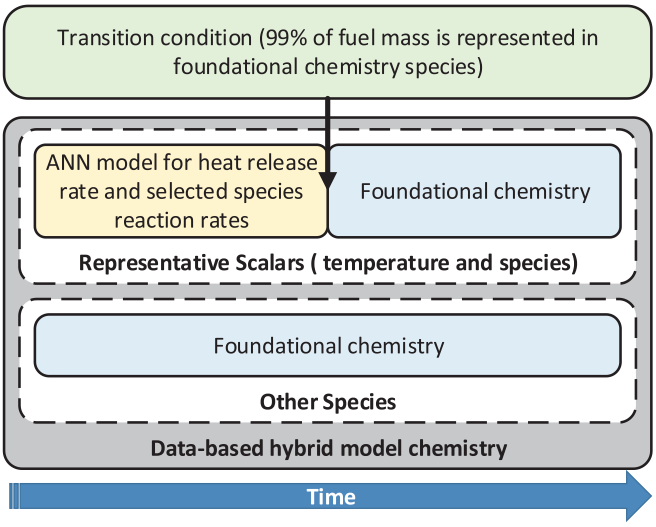

Fig. 2. Data-based hybrid model reaction rates calculation by ANN chemistry before transition to the foundational chemistry.

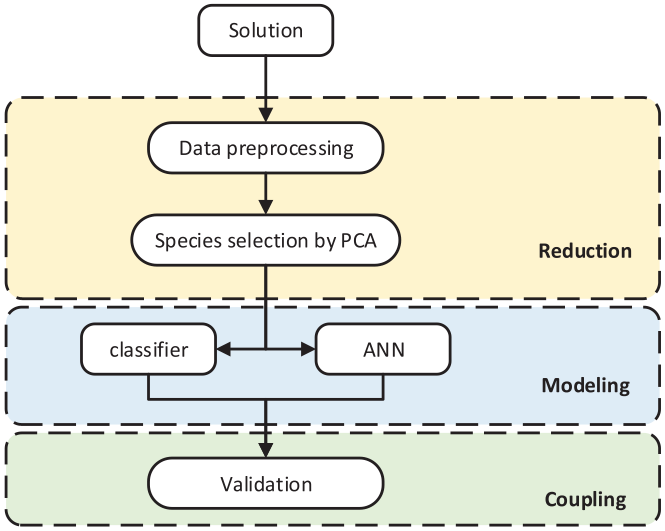

Fig. 3. Data-based hybrid model scheme.

the modeled species are the fragments, we model, using ANN, the reaction rates of representative species, which are used to track the evolution of the mixture from pyrolysis to the subsequent oxidation stages. These species, compared to fragments, tend to consist of simpler hydrocarbons and other oxidation species. Although they can be modeled using this foundational chemistry at later stages, they are modeled differently during pyrolysis stages to account for their interactions with larger fragments and the fuel.

Figure 2 illustrates the process of integration in time of all species in the mechanism for the case of using a C_0–C_4 or a C_0–C_2 foundational chemistry. The representative species are integrated during the initial stages of the chemistry integration using an ANN regression of the reaction rate built from numerical data of the evolution of the detailed mechanism. Subsequently, these species are integrated with the foundational chemistry. The remaining species (other than the representative species) are handled entirely by the foundational chemistry. The selection of the representative species out of the foundational chemistry species is carried out using *a priori* simulation data and principal component analysis [19] as discussed below. As a data-based framework, the proposed hybrid model is designed to be adaptive to a set of conditions dictated by the data on which the model is trained. This should not be perceived as a deficiency in the model; rather, it is an attribute of the model to enable a more aggressive reduction in the description of the chemistry if we know in advance the nature of the composition space accessed during simulations. Generalizing to a broader set of conditions requires a diverse set of data that is applicable to more scenarios for mixing and chemistry. From this perspective, the training for the hybrid model is not much different from that based on conventional chemistry reduction schemes. And, potentially, a hybrid chemistry model can be combined with a conventional chemistry reduction approach as a prior step before the development of the hybrid model.

The training steps to construct the chemistry acceleration model are summarized in Fig. 3. First, the model must be built on an existing solution that is representative of the problem under consideration, such as homogeneous chemistry for systems characterized by faster chemistry compared to transport, or canonical premixed/non-premixed flames or other stochastic 0D reactor models. This solution is built using a reference detailed or skeletal mechanism that we wish to reduce to accelerate chemistry integration. Of course, it can also be based on a HyChem hybrid chemistry models that couples global reaction steps for the fuel fragments with foundational chemistry. The second key step is the reduction step, which involves principal component analysis (PCA) [19] on the simulation data. PCA is used to select the representative species for which the reaction rates are modeled using an ANN regression.

The calculation of the representative species reaction rate during the pyrolysis stage and potentially part of the oxidation stage is carried out using ANN regression based on an input defining the mixture state using the representative species and temperature or a subset of this vector. At this step, criteria for determining the transition from the ANN regression model to the foundational chemistry model for the representative species' reaction rates are determined using a pattern recognition network (PRN) classifier discussed below. Finally, the validation step is carried out to compare the predictions of the hybrid model to results from detailed chemistry simulations. Once the hybrid model is validated, it can now be used in simulations as a chemistry acceleration scheme.

As indicated above, we use PCA for the representative species' selection. PCA is a feature/variable reduction approach that can reduce the representation of data from a set of correlated variables to a linear transform of uncorrelated variables. By this transformation, most of the data variance resides in the first principal components (PCs). PCA is becoming increasingly popular as a method for composition space reduction [20–28]. Its implementation in mechanism reduction is not new in the combustion literature [19,29–34]. However, it has been implemented primarily on variables that represent the sensitivities of thermo-chemical scalars or global measures of combustion (e.g. the flame speed) on reaction rate parameters to identify the key reactions to retain. Another novel implementation of PCA by D'Alessio et al. [35] implemented PCA as a method to partition the composition space into different clusters where different reduced mechanisms can be implemented. In the present study, we implement PCA on thermo-chemical scalars' reaction rates to identify the key species to retain.

PCA operates on existing data, which correspond in the present model to detailed chemistry simulations of a combustion problem. More specifically, in the present study, we consider a homogeneous system of an initial fuel-air mixture at high temperature. PCA takes the data from the temporal evolution of the temperature and species profiles until equilibrium is reached. It is implemented first on the full set of N species' reactions $\boldsymbol{\Omega} = (\omega_1, \omega_2, \ldots, \omega_N)$, excluding the fuel, in the detailed mechanism over time incre-

ments. The first step selects the first set of species; and a second PCA is implemented on the retained species from the first PCA. With a single-step PCA as a potential alternative, the species' selection may vary significantly by small changes in their criteria. This prospect is eliminated by the 2-step PCA hierarchical procedure.

To carry out the PCA, each species reaction rate is centered by subtracting its mean over all variable as follows:

$$\omega_i^* = \omega_i - \bar{\omega}. \tag{3}$$

The two-step PCA procedure is implemented as follows. For the first PCA, the constant matrix $\mathbf{Q}$ of eigenvectors of the symmetric covariance matrix $\mathbf{C}$ of size $N \times N$ is evaluated based on the solution of $\mathbf{\Omega}^*$ over the different time increments and initial conditions. Mathematically, components C_{ij} correspond to the covariance of the reaction rates of the ith and jth species, ω_i^* and ω_j^*, over all values of the solution and is expressed as follows:

$$C_{ij} = \frac{1}{M}\sum_{m=1}^{M} \omega_i^{*m}\omega_j^{*m}. \tag{4}$$

where M is the total number of discrete data points from the solution of $\mathbf{\Omega}^*$. The PCs' vector $\mathbf{\Psi}$ is related to the original thermo-chemical scalars' reaction rate vector as follows:

$$\mathbf{\Psi} = \mathbf{Q^T \Omega^*}. \tag{5}$$

At this stage, it is important to order the PCs based on the magnitude of the eigenvalues in a descending order. Every PC, $\mathbf{\Psi}_j$, also may be expressed as a linear combination of the centered thermo-chemical scalars' reaction rate vector:

$$\Psi_j = q_{1,j}\,\omega_1^* + \, q_{2,j}\,\omega_2^* + .. + q_{N,j}\,\omega_N^* \tag{6}$$

where the $q_{i,j}$ are the coefficients of matrix $\mathbf{Q^T}$ corresponding to thermo-chemical scalar $\mathbf{\Omega}_i^*$. The first few PCs represent most of the variance in the solution vector Ω. Therefore, we retain only the first N_{PC} PCs that contain a cumulative variance of 99%. These retained PCs can be expressed in terms of the normalized thermo-chemical scalars as follows:

$$\mathbf{\Psi_{red}} = \mathbf{A^T \Omega^*}. \tag{7}$$

The matrix A contains the leading N_{PC} eigenvectors of $\mathbf{Q}$. The same expression (7) is applied to relate the retained PCs to the species' reaction rates vector. However, j ranges from 1 to N_{PC} instead of from 1 to N.

At this stage, we reorder the sum in Eq. (6) such that the terms with the highest magnitudes of the coefficients $q_{i,j}$ are listed first in a descending order again. This reordering also identifies the importance of the associated thermo-chemical scalars. To identify the most important species, we implement a cut-off criterion as follows:

$$\frac{\sum_{i=1}^{b} |q_{i,j}|}{\sum_{i=1}^{N} |q_{i,j}|} \times 100 > \text{cutoff}\,\%\,, \quad j = 1, \ \ldots, \ N_{PC}. \tag{8}$$

The goal is to determine the number b of the reordered thermo-chemical scalars that represents a cumulative magnitude of the coefficients $q_{i,j}$ with the dominant contributions to the PCs. For example, in the present study of n-$C_{12}H_{26}$, n-C_7H_{16} and n-$C_{10}H_{22}$, we obtain a value for b ranging between 40 and 50 species, which yields a cutoff percentage of 99%. The identified species from this range include C_1-C_4 hydrocarbons along with a few more complex hydrocarbons, including C_5H_{10}, C_6H_{12} and C_7H_{14}.

The next PCA is carried out on the selected species from the first PCA. For this step, 3 conditions are imposed on the number of retained PCs: (1) these retained PCs must account for more than 99% (or similar threshold) of the data variance and (2) the cutoff condition (based on Eq. (8)) and (3) the selected species from the list must be a part of the foundational chemistry. In this study,

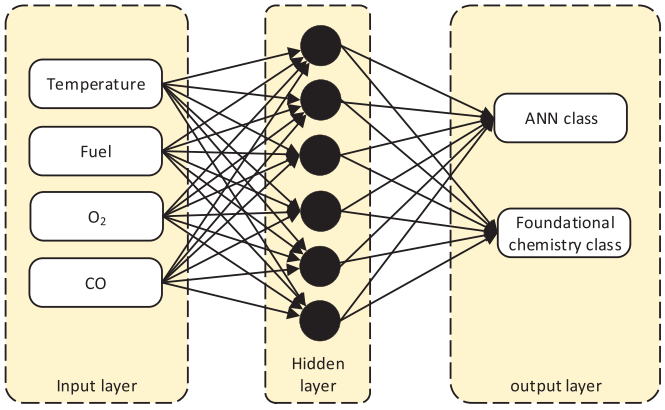

Fig. 4. A pattern recognition network (PRN) architecture of 4 nodes in the input layer, one hidden layer with 6 neurons, and two classes.

the cutoff condition percentage is investigated in the range of 80% to 99% and two foundational chemistries are used based on two widely-used mechanisms, USC Mech-II [36], a C_0–C_4 mechanism, and GRI-Mech 1.2 [37], a C_0–C_2 mechanism. Although this earlier version of the GRI mechanism is not as comprehensive as the later versions, especially as far as including more C_2-related steps, it remains a valid foundational chemistry to capture the oxidation phase of complex hydrocarbons. The number of selected species, N_r, from the second PCA ranges from 12 to 19 species for the 3 fuels considered.

The next step in the model development is an ANN-based regression based on a multi-layer perceptron (MLP) architecture for the reaction rate of the final selection of the species (the output) in terms of a subset of the representative species and temperature (the input). We have found that specifying the entire list of selected representative species for input in the ANN is not necessary and that a subset of these species is sufficient. In fact, the ANN in this study is established with only 3 species and the temperature as input. These species include the fuel, O_2, and CO. These scalars are adequate markers for the progress of reaction from pyrolysis to oxidation. Also, they are the same variables used to select the reaction rate model using PRN. For the ANN/PRN training and most importantly during *a posteriori* simulations, the selection of a subset of the representative species as an input for the ANN/PRN reduces the computational cost of training and selection of the chemistry model. Based on the present results, this subset recovers all representative species' reaction rates. Each representative species is modeled by a single ANN to independently optimize the ANN architecture for accuracy and computational efficiency.

The criterion for transition from ANN regression to foundational chemistry for the representative species is set during the training process based on the simulation data. For homogeneous chemistry, we set a criterion when 99% of the fuel C and H are represented by the species in the foundational chemistry. We use this criterion to train an ANN-based classifier, a pattern recognition network (PRN), to identify which model for chemistry, ANN-based or foundational chemistry, to use during *a posteriori* simulations. The PRN classifier may not be needed within the context of the homogeneous system considered in the present study. However, it will be important in more complex configuration when both mixing and reaction are present or when the model is trained on different initial mixtures of fuel and oxidizer.

Figure 4 shows the network architecture for a PRN. It includes an input layer, a single hidden layer, and an output layer. In principle, the entire set of transported thermo-chemical scalars or the entire list of representative species can be represented by neurons in the input layer. Again, we have found that a subset of

the thermo-chemical scalars' vector is sufficient to decide which chemistry model to use for the representative species based on the state of the mixture. In the present study, this subset consists of temperature, the fuel, O_2 and CO mass fractions. Also, for effective training, at least one hidden layer in the PRN architecture is needed. In the present study, we use one hidden layer with 6 neurons.

The network calculates a probability based on the input representation of the thermo-chemical state of whether the chemistry model falls into the ANN-based class or the foundational chemistry class. Based on the value of this probability, the decision is made to use either model based on its higher probability.

To illustrate the procedure for implementing the PRN as a classification algorithm, we briefly discuss the essential steps of training (on the training data) and implementing the network *in situ* during a simulation involving the hybrid model. In the beginning, the training data is labeled with either a value of 0 or 1, depending on whether a given state needs to be modeled with ANN chemistry or with the foundational chemistry models, respectively, based on an *a priori* criterion for classification. First, we select the PRN architecture, which in this case corresponds to 4 neurons at the input, one hidden layer with 6 neurons, and one output neuron, which will be labeled as belonging to one of two classes.

In a second step, the PRN is trained by matching the input training data with their labeled classes. This results in the determination of weights that measure the strengths of the connections from inputs to neurons in the hidden layers and these neurons to the output. Mathematically, the value of the output, which is interpreted as a probability, Prob, is expressed as:

$$\text{Prob} = f\left(\sum_{i=1}^{6} w_i^{(1)}\, a_i^{(1)} + b^{(1)}\right) \tag{9}$$

where f is the activation function, which is, in the present study, the sigmoid function. $w_i^{(1)}$ is the weight of the connection between the ith neuron in the hidden layer and the output, which measures the strength of this connection. $b^{(1)}$ is the bias neuron value in the hidden layer and $a_i^{(1)}$ is the value of the ith neuron in the hidden layer. This value, in turn, can be related to the values at the input using a similar expression to Eq. (9):

$$a_i^{(1)} = f\left(\sum_{j=1}^{4} w_{ji}^{(0)}\, a_j^{(0)} + b^{(0)}\right) \tag{10}$$

where $w_{ji}^{(0)}$ is the weight of the connection between the jth neuron in the input layer and the ith neuron in the hidden layer. $a_j^{(0)}$ and $b^{(0)}$ are the values of the jth neuron and the bias neuron in the input layer. The bias neurons serve to shift the activation function f, hence enabling more flexibility in training the network.

Once the training is completed and the connections' weights are determined, the network is ready to be implemented *in situ* during a simulation using the hybrid model. Given a state of the solution, a subset of the solution vector represented by the 4 inputs is used to determine the probability according to Eqs. (9) and (10). These probabilities, then, are converted to 0's or 1's, depending to their relative proximity to these 2 limits, thus identifying the class of reaction models to use for the state.

It is important to emphasize the importance of developing *a priori* criteria for the "classification" of the reaction model for the representative species. With a single set of training conditions, it is easier to set a threshold criterion for the transition from a reaction to another. However, if the model data is to be derived from a broad range of simulations (e.g. a combination of reactor models or under non-homogeneous conditions), the criteria for transition can be set during the training stage using a PRN and the corresponding network can provide an unbiased classification of the model during *a posteriori* simulations.

Table 1
Number of species in detailed mechanism for n-dodecane, n-heptane, and n-decane.

Fuel	n-dodecane	n-heptane	n-decane
Detailed mechanism	JetSurF2.0 [38]	Mehl et al. [39]	Dooley et al. [40]
Total species	348	654	1599
Species with $> C_4$	250	381	1207
Species with C_3–C_4	57	200	313
Species with C_0–C_2	41	73	79

The feedforward network for ANN-based regression shares common elements of its architecture with the PRN architecture shown in Fig. 4. The input layer is identical to that of the PRN architecture and uses only 4 inputs. Given that the data is relatively simple, also only one hidden layer is used with the number of neurons varying from 5 to 15 depending on the complexity of the reaction rate to be modeled. Training for a broader range of mixture conditions may require a deeper network to accommodate the range of mixing and reaction interactions. Moreover, the training is done individually for each representative species reaction rate. Therefore, the output has only one neuron and different networks, one for each representative species, are used. The transfer function connecting a given neuron to neurons from a previous layer is the hyperbolic tangent sigmoid transfer function. The ANN weights and biases are updated during the training process by minimizing the loss function by a Levenberg-Marquardt backpropagation algorithm. The loss function is chosen as the mean squared error (MSE). The training data used to build the hybrid model is divided randomly into three sets, a learning set, a validation set, and a test set with proportions of 60/20/20%, respectively, and fed to the training algorithm. The training times with 3000 epochs/iterations for each reaction rate takes less than 10 min on a single CPU workstation.

It is important to note that there is significant room for optimization of the network architecture, including its depth and number of neurons in each hidden layer, as well as whether to train all representative species' reaction rates individually or in groups. This optimization will be attempted in the future to accommodate larger datasets with more complex composition spaces.

Finally, in *a posteriori* simulations, the representative species are integrated with the remaining species of the foundational chemistry. The solutions for the temperature and a subset of the representative species is used to distinguish the regime for modeling the representative scalars reaction rates using the PRN network. The same set of inputs is used to determine the representative species and temperature chemical source terms in the governing equations if the ANN model is needed. In addition to the input variables for the networks, the weights associated with each network and its architecture are used. As described above, these weights are determined during the modeling step (see Fig. 3).

3. Results and discussion

3.1. Run conditions

The present model is validated using a simple reactor model based on homogeneous chemistry at constant volume for 3 fuels: n-dodecane, n-heptane and n-decane. The characteristics of the detailed mechanisms for these fuels are summarized in Table 1. The JetSurF2.0 mechanism is used as a detailed chemistry mechanism for the n-dodecane. It consists of 348 species and 2163 reactions [38]. For n-heptane, a detailed mechanism by Mehl et al. [39] is used. It consists of 654 species and 2827 reactions. A jet fuel surrogate mechanism by Dooley et al. [40] is used for n-decane. It

Table 2
List of representative species for n-dodecane, n-heptane, and n-decane.

Fuel	n-$C_{12}H_{26}$		n-$C_{10}H_{22}$		n-C_7H_{16}	
Hybrid model	Foundational chemistry used with the hybrid model					
	USC- Mech	GRI- Mech	USC- Mech	GRI- Mech	USC- Mech	GRI- Mech
common species			O_2 CO_2 CO H_2O H_2 OH H O CH_3 CH_4 C_2H_2 C_2H_4			
	C_2H_6 C_3H_6 C_4H_8-1 HCCO	C_2H_6 HCCO	C_2H_5 C_3H_6 C_4H_8-1 CH_2O	CH_2O	C_2H_6 C_3H_6 C_4H_8-1	C_2H_6

consists of 1599 species and 6633 reactions. Of course, it is possible to select a subset of the reactions and species in each mechanism or based on the detailed chemistry calculations on skeletal mechanism. However, one of our principal goals is to investigate whether the present model can describe the chemistry of these fuels by using a smaller foundational chemistry mechanism, either USC Mech-II [36] with 111 species and 784 reactions or GRI-Mech 1.2 with 32 species and 177 reactions [37]. Based on the information in Table 1, a large portion of the detailed mechanisms is dedicated to the chemistry of species larger than C_4 and an important portion is represented by C_3–C_4 species. Therefore, by eliminating the complex hydrocarbons and their associated reactions, potential computational saving and acceleration can be achieved. Further chemistry acceleration can be achieved by a reduction of the foundational chemistry. Standard approaches for chemistry reduction can be coupled with our present hybrid chemistry model.

The governing energy and species equations for the homogeneous reactions are:

$$\frac{dT}{dt} = -\frac{\sum_{l=1}^{N} u_l \, \dot{\omega}_l \hat{W}_l}{\rho \bar{c}_v} \, , \tag{11}$$

$$\frac{dC_k}{dt} = \dot{\omega}_k \, , \; k = 1 \text{ to } N \tag{12}$$

where T is the temperature and $u_l, \dot{\omega}_l$ and $\hat{W}_l$ are the lth species internal energy, reaction rate and molecular weight, respectively, ρ is mixture density, $\bar{c}_v$ is mixture specific heat at constant volume, and C_k is the kth species concentration. In the above equations, N corresponds to number of species in the foundational chemistry mechanism. The governing equations (11) and (12) are integrated using Senkin [41].

In this study, we train the model for an initial pressure of one atmosphere, a stoichiometric composition of fuel and air and vary the initial temperature from 1400 K to 1500 K at 10 degrees increments. For validation, we also investigate the ability of the model to "interpolate" by investigating initial temperatures at 1425 and 1475 K.

Our present results show that the combination of 4 PCs and a cutoff percentage of 90% yield the best performance in term of solution accuracy and ANN training time for all studied fuels. The representative species for all studied fuels are listed in Table 2. The data-based hybrid model predicts a common set of representative species among all studied fuels with both foundational chemistries. They include: O_2, CO_2, CO, H_2O, H_2, OH, C_2H_4, H, C_2H_2, O, CH_3, and CH_4.

3.2. Model performance

In this section, we present comparisons of the predictions of the data-based hybrid chemistry model with the detailed chemistry results for the 3 fuels considered and the two foundational chemistries adopted. Figures 5 and 6 show temporal profiles' comparisons among the data-based hybrid model coupled with USC Mech-II, data-based hybrid model coupled with GRI-Mech 1.2 and the detailed mechanism at 1425 K and 1475 K, respectively, initial temperatures for n-dodecane. The results are presented for temperature and select species mole fractions that are either modeled with the hybrid chemistry approach or the foundational chemistry. Here, the modeled species by ANN in the hybrid model for n-dodecane combined with USC Mech-II are O_2, CO_2, CO, H_2O, H_2, OH, C_2H_4, H, C_2H_2, O, CH_3, CH_4, HCCO, C_2H_6, C_4H_8-1 and C_3H_6. Therefore, in the figures, HO_2, CH_3OH, CH_3O and H_2O_2 are modeled entirely by the foundational chemistry. The results show that the data-based model can predict the evolution of a subset of these species as well as the ones modeled initially by ANN and temperature accurately. Nonetheless, the present discrepancies still capture the trends of the temporal profiles. The resulting computational saving is approximately 5 times; although, there was no attempt to optimize the performance of the model.

Next, we substitute the foundational chemistry by using the simpler GRI-Mech 1.2 instead of the USC Mech-II. The modeled species by ANN with the GRI-Mech 1.2 include: O_2, CO_2, CO, H_2O, H_2, OH, C_2H_4, H, C_2H_2, O, CH_3, CH_4, HCCO, and C_2H_6. Therefore, 4 of the species shown in Figs. 5 and 6 are modeled with the foundational chemistry. The use of the simpler GRI-Mech 1.2 pushes the transition point for the hybrid reaction rate model to a later time. This is expected since a transition to a dominant C_0–C_2 chemistry occurs later than the transition to a dominant C_0–C_4 chemistry. The figure shows an excellent agreement for the temperature and a number of major species modeled by ANN, reasonable predictions for HO_2 and CH_3O, which are modeled by the GRI-Mech 1.2 and some discrepancies for CH_3OH and H_2O_2. Again, despite the presence of some quantitative discrepancies, the trends of the temporal profiles are well captured by the model and the final equilibrium conditions are reasonably matched with the GRI-Mech 1.2 as the foundational chemistry model. The coupling of the model with the

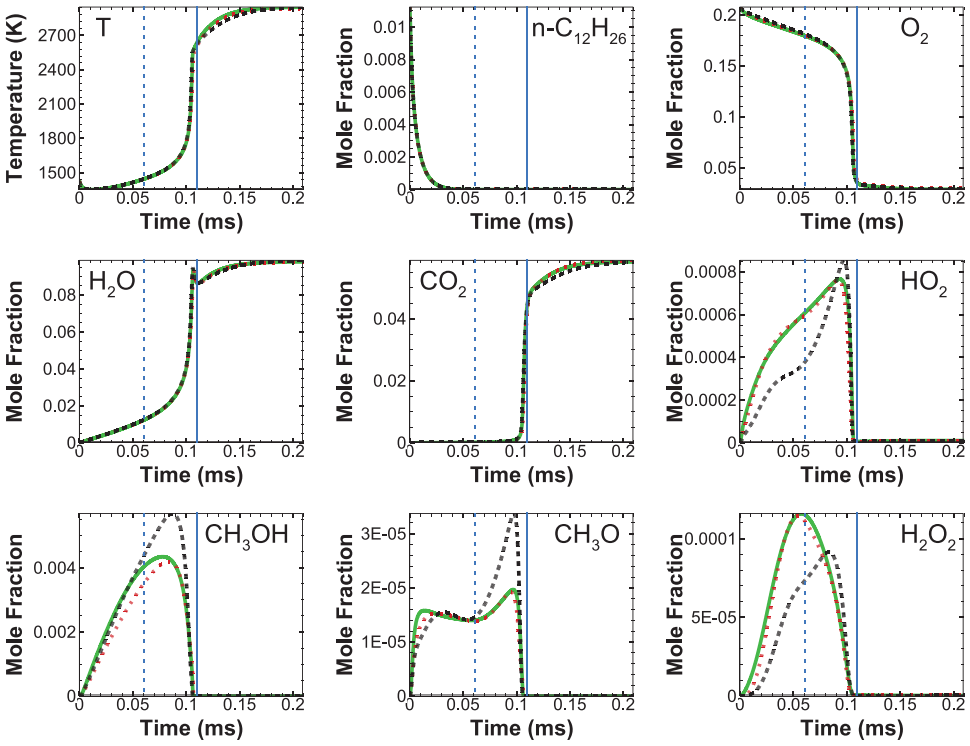

Fig. 5. Stoichiometric n-dodecane/air at 1425 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical dashed blue line and the vertical solid blue line represent the transitions between the ANN regression model and the foundational chemistry based on the USC Mech-II and the GRI-Mech 1.2, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

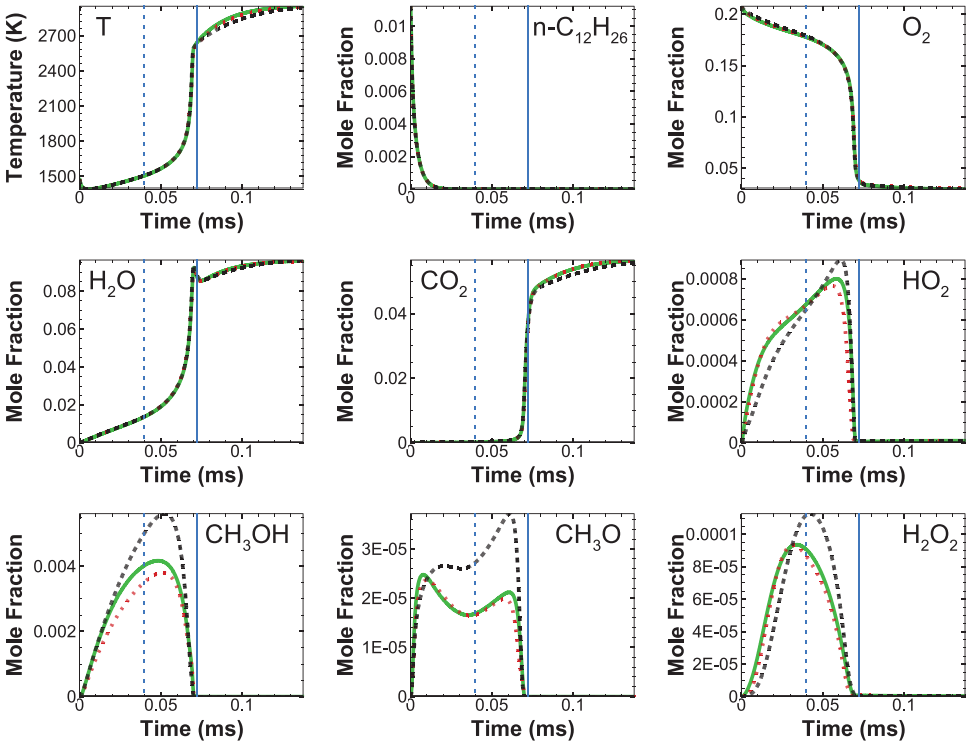

Fig. 6. Stoichiometric n-dodecane/air at 1475 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical blue lines correspond to the same criteria as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

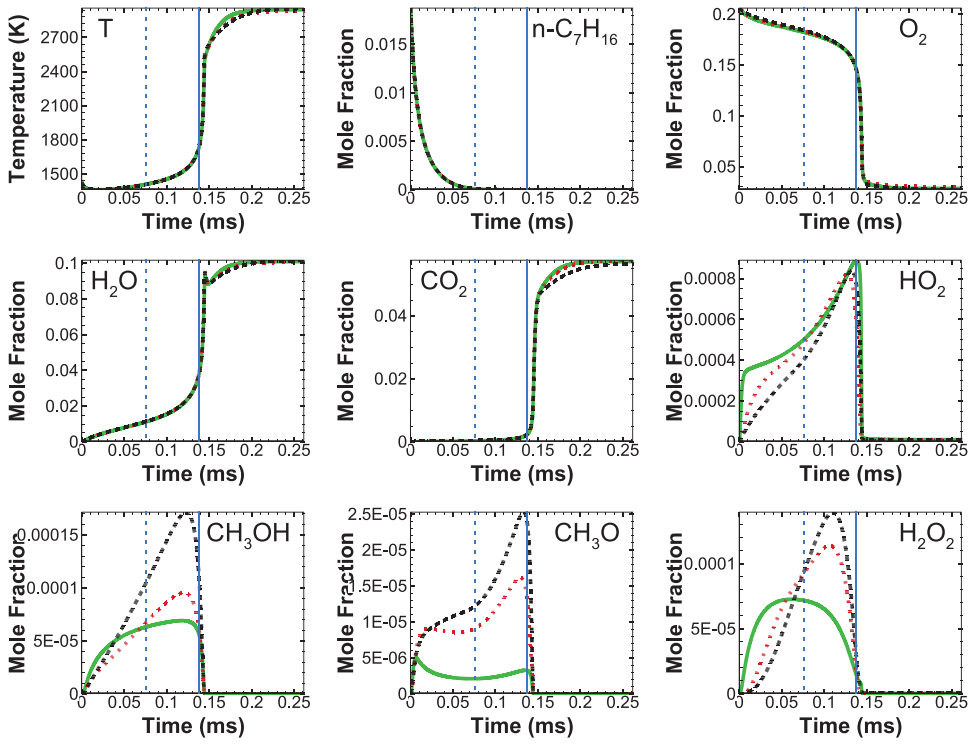

Fig. 7. Stoichiometric n-heptane/air at 1425 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical blue lines correspond to the same criteria as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

GRI-Mech 1.2 is achieved with a computational saving of an order of magnitude.

Next, we carry out comparisons based on n-heptane simulations with the data-based hybrid model. Figures 7 and 8 mirror the conditions in Figs. 5 and 6, except that the fuel now is n-heptane. Moreover, the list of modeled species with ANN is different for this fuel. With the USC Mech-II as the foundational chemistry model, the following species are selected using the PCA procedure outlined above: O_2, CO_2, CO, H_2O, H_2, OH, C_2H_4, H, C_2H_2, O, CH_3, CH_4, C_2H_6, C_4H_8-1and C_3H_6. Again, in Figs. 7 and 8, HO_2, CH_3OH, CH_3O and H_2O_2 are modeled with the foundational chemistry.

The following list is adopted for coupling with the GRI-Mech 1.2: O_2, CO_2, CO, H_2O, H_2, OH, C_2H_4, H, C_2H_2, O, CH_3, CH_4, and C_2H_6. Figures 7 and 8 confirm essentially the same observations made for n-dodecane. The use of the USC Mech-II as foundational chemistry yields better predictions than the model simulations based on the GRI-Mech 1.2 for foundational chemistry. The computational saving with the model for n-heptane is one order of magnitude with USC Mech-II and two orders of magnitude with the GRI-Mech 1.2.

The same trends are established for similar comparisons between model predictions and detailed chemistry calculations for n-decane, which are shown in Figs. 9 and 10. For this fuel, the following species are selected for ANN modeling with the USC Mech-II: O_2, CO_2, CO, H_2O, H_2, OH, CH_4, C_2H_4, H, C_2H_2, O, CH_3, CH_2O, C_4H_8-1 and C_3H_6; while a different set as listed in Table 2 is used with the GRI-Mech 1.2. The computational saving with the model for n-decane is even higher compared to the other fuels, yielding 3 orders of magnitude in acceleration.

The above results demonstrate that a hybrid chemistry model that is based on an ANN-based reaction model for a set of representative scalars (species and temperature) coupled with simpler foundational chemistry can be a viable chemistry acceleration scheme. The model performs slightly better with the C_0–C_4 foundational chemistry; but we can also see that the simpler foundational chemistry also can yield comparable or better results for some species in the n-decane simulations. As suggested in Fig. 1, the accuracy of the model may be attributed primarily to the value of the assumption associated with the transition as well as the role played by modeled representative scalars in the reaction of other species in the foundational chemistry.

At this stage, it is useful to reiterate that the hybrid approach presented in Ref. [17] and the present approach are based on tracking the pyrolysis stage during high-temperature oxidation (HTO) using a set of representative species. Both approaches yield very good comparisons with detailed mechanism results. In Ref. [17], as proposed in the HyChem approach [10–16], a common set (among a wide range of complex fuels) of pyrolysis fragments represent a natural choice for representative species during HTO. The present approach develops a more systematic strategy for the selection of these species using PCA. This strategy has been recently implemented to derive similar representative species during low-temperature fuel oxidation (LTO) [42]. In contrast with HTO with its common fragments, the LTO process is initiated through the formation of fuel-specific intermediates [43]. The present strategy enabled the selection of representative species that are simpler than these intermediates; and yet they can track even the early stages of LTO.

4. Conclusions

A data-based model for chemistry acceleration to integrate the chemistry of complex fuels at high temperature is developed. The model is based on a hybrid scheme that involves data-based models for representative species combined with the solution of the remaining species using a foundational chemistry mechanism. The selection process for the representative species is carried out using a 2-step PCA on the reaction rate data from detailed simulations.

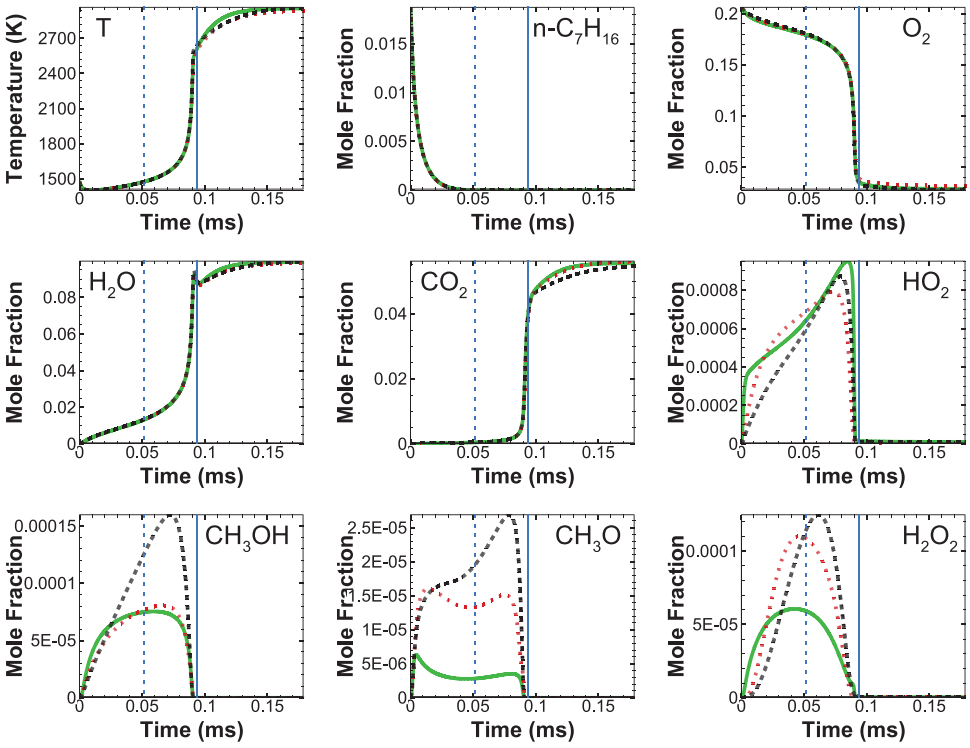

Fig. 8. Stoichiometric n-heptane/air at 1475 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical blue lines correspond to the same criteria as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

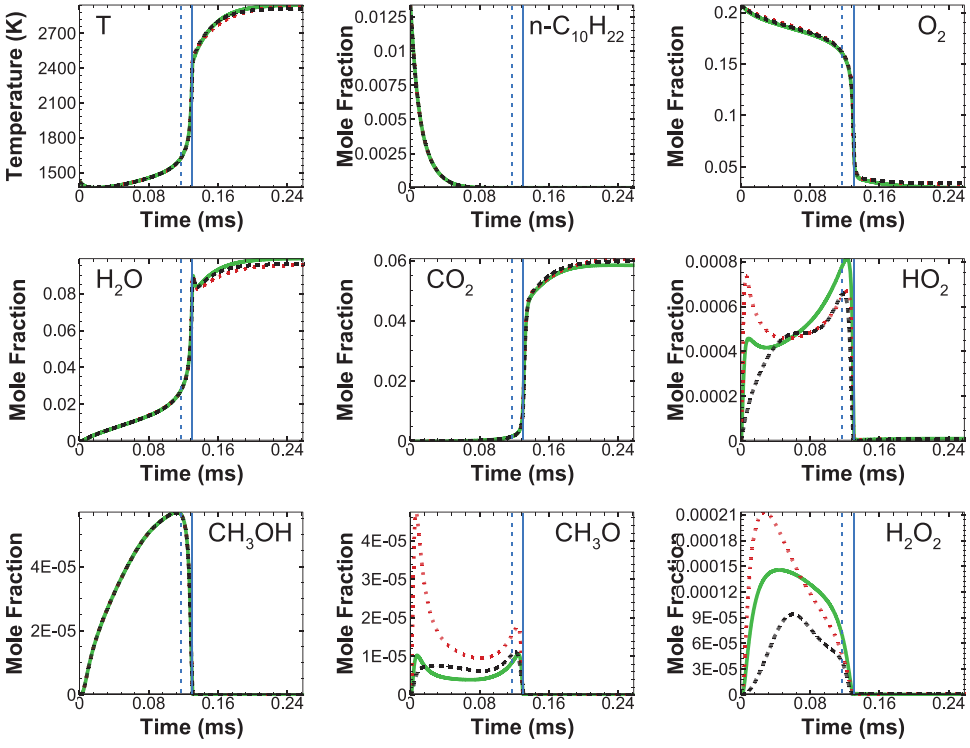

Fig. 9. Stoichiometric n-decane/air at 1425 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical blue lines correspond to the same criteria as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

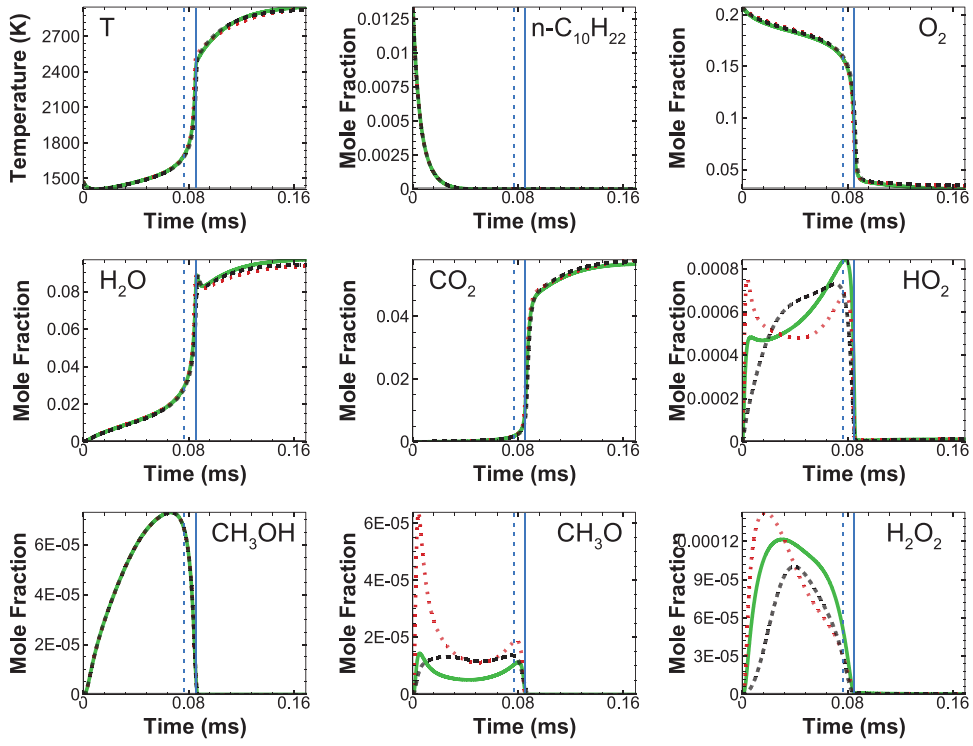

Fig. 10. Stoichiometric n-decane/air at 1475 K. Temporal profiles' comparisons of the hybrid model (dashed red: USC Mech-II, dashed black: GRI-Mech 1.2) and the detailed mechanism (solid green) predictions. The vertical blue lines correspond to the same criteria as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Once the selection of the representative species is completed, their reactions rates are modeled using ANN starting with the available computational data. These species' modeled reactions account for their interactions with the complex hydrocarbons. This ANN-based reaction rate model is coupled in time with the solution of the remaining species using a smaller foundational chemistry. The criteria for transition from the model to the foundational chemistry are implemented using a PRN-based classification approach. This approach is implemented during the training stage using the simulation data and a reduced input.

The validation of the hybrid model by comparison with detailed chemistry calculations for 3 fuels, n-dodecane, n-heptane and n-decane, shows that the model can yield reasonable results by adopting either C_0–C_2 or C_0–C_4 foundational chemistries and can result in acceleration of the chemistry integration process from 1 to 3 orders of magnitude. The reduction process also sifts through the complexity of the detailed mechanism to identify the key species to model and transport.

The model is built on adapting the chemistry for a specific class of problems. It can be generated based on data from surrogate problems that spans the same composition space of the configuration of interest. Such problems include 0D reactor models, such as perfectly stirred reactors, 1D flame models or low-dimensional stochastic models, such as the linear-eddy model (LEM) or the one-dimensional turbulence (ODT) model. The generation of data from these models is inherently cheaper than the detailed multi-dimensional reacting flow simulations of interest. In these simulations, accounting for chemistry often makes up the lion's share of the computational cost.

Finally, the choice of the surrogate model is critical to the success of the implementation of the hybrid chemistry model proposed here. This choice can determine how accurate these predictions are. However, an equally important consideration is that the accessed composition space during multi-dimensional simulations must not exceed the bounds of the corresponding composition space of the surrogate models. Criteria should be implemented to identify such events. However, and most importantly, strategies must be implemented to predict conditions outside the bounds of the hybrid chemistry model. Such events can be handled by accessing a different model, such as a more detailed chemical description, or by retraining *in situ* the hybrid chemistry approach to accommodate the new events.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The first author would like to acknowledge the support of King Khalid University in Abha, Saudi Arabia.

References

[1] T. Turanyi, A.S. Tomlin, Reduction of reaction mechanisms, in Analysis of Kinetic Reaction Mechanisms, Springer (2014), pp. 183–312.
[2] S.B. Pope, Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation, Combust. Theor. Model. 1 (1997) 41–63.
[3] S.R. Tonse, N.W. Moriarty, M. Frenklach, N.J. Brown, Computational economy improvements in PRISM, Int. J. Chem. Kinet. 35 (2003) 438–452.
[4] U. Maas, S.B. Pope, Simplifying chemical kinetics – intrinsic low-dimensional manifolds in composition space, Combust. Flame 88 (1992) 239–264.
[5] S.H. Lam, D.A. Goussis, The CSP method for simplifying kinetics, Int. J. Chem. Kinet. 26 (1994) 461–486.
[6] L. Liang, J.G. Stevens, S. Raman, J.T. Farrell, The use of dynamic adaptive chemistry in combustion simulation of gasoline surrogate fuels, Combust. Flame 156 (2009) 1493–1502.
[7] F. Continuo, H. Jeanmart, T. Lucchini, G. D'Errico, Coupling of in situ adaptive tabulation and dynamic adaptive chemistry: an effective method for solving combustion in engine simulations, Proc. Combust. Inst. 33 (2011) 3057–3064.

[8] W. Sun, Y. Ju, A multi-timescale and correlated dynamic adaptive chemistry and transport (CO-DACT) method for computationally efficient modeling of jet fuel combustion with detailed chemistry and transport, Combust. Flame 184 (2017) 297–311.
[9] X. You, F.N. Egolfopoulos, H. Wang, Detailed and simplified kinetic models for *n*-dodecane oxidation: the role of fuel cracking in aliphatic hydrocarbon combustion, Proc. Combust. Inst. 32 (2009) 403–410.
[10] R. Xu, H. Wang, D.F. Davidson, R.K. Hanson, C.T. Bowman, F.N. Egolfopoulos, Evidence supporting a simplified approach to modeling high-temperature combustion chemistry, 10th U.S. National Combustion Institute Meeting, The Combustion Institute, College Park, MD, 2017 April 23–26.
[11] R. Xu, D. Chen, K. Wang, T. Tao, J.K. Shao, T. Parise, Y. Zhu, S. Wang, R. Zhao, D.J. Lee, F.N. Egolfopoulos, D.F. Davidson, R.K. Hanson, C.T. Bowman, H. Wang, HyChem Model: application to petroleum-derived jet fuels, 10th U.S. National Combustion Institute Meeting, The Combustion Institute, 2017 April 23–26.
[12] K. Wang, R. Xu, T. Parise, J.K. Shao, D.J. Lee, A. Movaghar, D.F. Davidson, R.K. Hanson, H. Wang, C.T. Bowman, F.N. Egolfopoulos, Combustion kinetics of conventional and alternative jet fuels using a hybrid chemistry (HyChem) approach, 10th U.S. National Combustion Institute Meeting, The Combustion Institute, College Park, MD, 2017 April 23–26.
[13] K. Wang, R. Xu, T. Parise, J.K. Shao, D.F. Davidson, R.K. Hanson, H. Wang, C.T. Bowman, Evaluation of a hybrid chemistry approach for combustion blended petroleum and bio-derived jet fuels, 10th U.S. National Combustion Institute Meeting, The Combustion Institute, College Park, MD, 2017 April 23–26.
[14] R. Xu, D. Chen, K. Wang, H. Wang, A comparative study of combustion chemistry of conventional and alternative jet fuels with hybrid chemistry approach, 55th AIAA Aerospace Sciences Meeting, AIAA paper 2017-0607, 2017 January 9–13.
[15] H. Wang, R. Xu, K. Wang, C.T. Bowman, R.K. Hanson, D.F. Davidson, K. Brezinsky, F.N. Egolfopoulos, A physics-based approach to modeling real-fuel combustion chemistry-I. Evidence from experiments, and thermodynamic, chemical kinetic and statistical considerations, Combust. Flame 193 (2018) 502–519.
[16] R. Xu, K. Wang, S. Banerjee, J.K. Shao, T. Parise, Y.Y. Zhu, S.K. Wang, A. Movaghar, D.J. Lee, R.H. Zhao, X. Han, Y. Gao, T.F. Lu, K. Brezinsky, F.N. Egolfopoulos, D.F. Davidson, R.K. Hanson, C.T. Bowman, H. Wang, A physics-based approach to modeling real-fuel combustion chemistry–II. Reaction kinetic models of jet and rocket fuels, Combust. Flame 193 (2018) 520–537.
[17] R. Ranade, S. Alqahtani, A. Farooq, T. Echekki, An ANN based hybrid chemistry framework for complex fuels, Fuel 241 (2019) 625–636.
[18] R. Ranade, S. Alqahtani, A. Farooq, T. Echekki, An extended hybrid chemistry framework for complex hydrocarbon fuels, Fuel 251 (2019) 276–284.
[19] S. Vajda, P. Valko, T. Turányi, Principal component analysis of kinetic models, Int. J. Chem. Kinet. 17 (1985) 55–81.
[20] C.E. Frouzakis, Y.G. Kevrekidis, J. Lee, K. Boulouchos, A.A. Alonso, Proper orthogonal decomposition of direct numerical simulation data: data reduction and observer construction, Proc. Combust. Inst. 28 (2000) 75–81.
[21] S.J. Danby, T. Echekki, Proper orthogonal decomposition analysis of autoignition simulation data of nonhomogeneous hydrogen–air mixtures, Combust. Flame 144 (2006) 126–138.
[22] J.C. Sutherland, A. Parente, Combustion modeling using principal component analysis, Proc. Combust. Inst. 32 (2009) 1563–1570.
[23] A. Parente, J.C. Sutherland, B.B. Dally, L. Tognotti, P.J. Smith, Investigation of the MILD combustion regime via principal component analysis, Proc. Combust. Inst. 33 (2011) 3333–3341.
[24] H. Mirgolbabaei, T. Echekki, Nonlinear reduction of combustion composition space with kernel principal component analysis, Combust. Flame 161 (2014) 118–126.
[25] H. Mirgolbabaei, T. Echekki, The reconstruction of thermo-chemical scalars in combustion from a reduced set of their principal components, Combust. Flame 162 (2015) 1650–1652.
[26] O. Owoyele, T. Echekki, Toward computationally efficient combustion DNS with complex fuels via principal component transport, Combust. Theor. Model 21 (2017) 770–798.
[27] R. Ranade, T. Echekki, A framework for data-based turbulent combustion closure: a priori validation, Combust. Flame 206 (2019) 490–505.
[28] R. Ranade, T. Echekki, A framework for data-based turbulent combustion closure: a posteriori validation, Combust. Flame 210 (2019) 279–291.
[29] S. Vajda, H. Rabitz, R.A. Yetter, Effects of thermal coupling and diffusion on the mechanism of H2 oxidation in steady premixed laminar flames, Combust. Flame 82 (1990) 270–297.
[30] N.J. Brown, G.P. Li, M.L. Koszykowski, Mechanism reduction via principal component analysis, Int. J. Chem. Kinet. 29 (1997) 393–414.
[31] I.G. Zsély, T. Turányi, The influence of thermal coupling and diffusion on the importance of reactions: the case study of hydrogen-air combustion, Phys. Chem. Chem. Phys. 5 (2003) 3622–3631.
[32] T. Nagy, T. Turányi, Reduction of very large reaction mechanisms using methods based on simulation error minimization, Combust. Flame 156 (2009) 417–428.
[33] G. Esposito, H.K. Chelliah, Skeletal reactions based on principal component analysis: application to ethylene-air ignition, propagation, and extinction phenomena, Combust. Flame 158 (2011) 477–489.
[34] G. Esposito, M.J. Rahimi, H.K. Chelliah, P.D. Vogel, J.R. Edwards, Assessment of chemical kinetic modeling for silane/hydrogen mixtures in hypersonic applications, AIAA J 52 (2014) 2213–2222.
[35] G. D'Alessio, A. Parente, A. Stagni, A. Cuoci, Adaptive chemistry via pre-partitioning of composition space and mechanism reduction, Combust. Flame 211 (2020) 68–82.
[36] H. Wang, X. You, A.V. Joshi, S.G. Davis, A. Laskin, F. Egolfopoulos, C.K. Law, USC mech version II. High-temperature combustion reaction model of H2/CO/C1-C4 compounds. http://ignis.usc.edu/USC_Mech_II.htm, 2007.
[37] M. Frenklach, H. Wang, C.-L. Yu, M. Goldenberg, C.T. Bowman, R.K. Hanson, D.F. Davidson, E.J. Chang, G.P. Smith, D.M. Golden, W.C. Gardiner, V. Lissianski, http://www.me.berkeley.edu/gri_mech/.
[38] H. Wang, E. Dames, B. Sirjean, D. A. Sheen, R. Tango, A. Violi, J. Y. W. Lai, F. N. Egolfopoulos, D. F. Davidson, R. K. Hanson, C. T. Bowman, C. K. Law, W. Tsang, N. P. Cernansky, D. L. Miller, R. P. Lindstedt, A high-temperature chemical kinetic model of *n*-alkane (up to *n*-dodecane), cyclohexane, and methyl-, ethyl-, *n*-propyl and *n*-butyl-cyclohexane oxidation at high temperatures, JetSurF version 2.0, September 19, 2010 (http://web.stanford.edu/group/haiwanglab/JetSurF/JetSurF2.0/index.html).
[39] M. Mehl, W.J. Pitz, C.K. Westbrook, H.J. Curran, Kinetic modeling of gasoline surrogate components and mixtures under engine conditions, Proc. Combust. Inst. 33 (2011) 193–200.
[40] S. Dooley, S.H. Won, M. Chaos, J. Heyne, Y. Ju, F.L. Dryer, K. Kumar, C.-J. Sung, H. Wang, M.A. Oehlschlaeger, R.J. Santoro, T.A. Litzinger, A jet fuel surrogate formulated by real fuel properties, Combust. Flame 157 (2010) 2333–2339.
[41] A.E. Lutz, R.J. Kee, J.A. Miller, SENKIN: A FORTRAN Program for Predicting Homogeneous Gas Phase Chemical Kinetics with Sensitivity Analysis. No. SAND-87-8248, Sandia National Labs. Livermore, CAUSA, 1988.
[42] S.S. Alqahtani, Machine learning methods for chemistry reduction in combustion, Ph.D. Dissertation, Department of Mechanical and Aerospace Engineering, North Carolina State University, 2020.
[43] J. Zádor, C.A. Taatjes, R.X. Fernandes, Kinetics of elementary reactions in low-temperature autoignition chemistry, Prog. Energy Combust. Sci. 37 (2011) 371–421.