A functional-group-based approach to modeling real-fuel combustion chemistry -

I: Prediction of stoichiometric parameters for lumped pyrolysis reactions

Xiaoyuan Zhang^{*}, Kiran K. Yalamanchi, S. Mani Sarathy^{*}

*Clean Combustion Research Center, King Abdullah University of Science and Technology, Thuwal,
23955-6900, Saudi Arabia*

Abstract: Real fuels are complex mixtures of hundreds of molecules, which makes it challenging to unravel their combustion chemistry. Several approaches in the literature have helped to clarify fuel combustion, including multi-component surrogates, lumped fuel chemistry modeling, and functional-group based methods. This work presents an innovative advancement to the lumped fuel chemistry modeling approach, using functional groups for mechanism development (FGMech). Stoichiometric parameters of lumped fuel decomposition reactions dictate the population of the key pyrolysis products, previously obtained by fitting experimental data of real-fuel pyrolysis. In this work, a functional group-based approach is proposed, which can account for real-fuel variability and predict stoichiometric parameters without experimentation. A database of the stoichiometric parameters and/or yields of key pyrolysis products was first constructed for approximately 50 neat fuels, based on previous pyrolysis data and a lumped reaction mechanism we developed. The effects of functional groups on the stoichiometric parameters and/or yields of key pyrolysis products were then identified and quantified. A quantitative structure-yield relationship was developed by multiple linear regression (MLR) model, which was used to predict the stoichiometric parameters and/or yields of key pyrolysis products based on ten input features (eight functional groups, molecular weight, and branching index). Products from the pyrolysis of surrogate mixtures and real-fuels were predicted using the MLR model and validated against experimental data in the literature. Comparison with the stoichiometric parameters from the

^{*} Corresponding author. E-mail: xiaoyuan.zhang@kaust.edu.sa.

^{*} Corresponding author. E-mail: mani.sarathy@kaust.edu.sa.

HyChem experiment-based approach [Xu *et al.* Combust. Flame 193 (2018) 520-537] showed that the predicted values in this work were in reasonable agreement (generally within a factor of two). When the stoichiometric parameters in the jet fuel (POSF 10325) HyChem kinetic model were replaced with this functional-group based prediction, only minor discrepancies were observed in the predictions of key pyrolysis products and global combustion parameters (such as ignition delay times and laminar flame speeds). Sensitivity analysis on stoichiometric parameters revealed their different roles in predicting speciation and global parameters. The functional group approach for predicting stoichiometric parameters in this work was the first step towards developing FGMech for modeling real-fuel combustion chemistry. Further development of the FGMech model's thermodynamic, kinetic, and transport data will be presented in a following study.

Keywords: real-fuels; stoichiometric parameters; FGMech; functional groups; data research

1. Introduction

The prediction of combustion behavior in real-fuels—complex mixtures of hundreds of hydrocarbon molecules—is a major advancement in current kinetic modeling research [1-3]. In addition to the composite complexity of real-fuel mixtures, the size of chemical kinetic models is increased by considering large hydrocarbon compounds. Fortunately, in terms of high-temperature chemistry, it is the amount and nature of the small intermediate species population that dictate critical combustion parameters, like heat release, ignition and flame propagation [4,5]. Ranzi and coworkers were pioneers in applying lumped reactions to predict smaller intermediates formed in high temperature fuel oxidation, using the MAMOX++ program [6] to determine stoichiometric coefficients in lumped decomposition reactions [7]. Recently, a physics-based approach, HyChem [5], was proposed to model real-fuel combustion chemistry using lumped reaction schemes. Based on their study [5], it is found that the high-temperature combustion of real-fuels effectively occurs in two

separate stages: fuel pyrolysis (or oxidative pyrolysis) first to produce small critical intermediates, followed by the oxidation of these intermediates. Thus, obtaining the population of key pyrolysis products is the first step in the development of a lumped reaction model, whether for individual components or complex fuel mixtures.

For the sake of brevity, we briefly summarize key aspects of the HyChem approach as they pertain to this study. Readers are referred to the original HyChem papers [5,8] for a complete description of the methodology. The HyChem approach presents a comprehensive physics-based approach for modeling real-fuel combustion chemistry using a combination of stoichiometric parameters in lumped reactions and corresponding kinetic, thermodynamic, and transport parameters. In this work, we focus only on the stoichiometric parameters of lumped reactions. In the HyChem approach [5,8], H, CH₃, CH₄, C₂H₄, C₃H₆, 1-C₄H₈, IC₄H₈, benzene (C₆H₆) and toluene (C₇H₈) are identified as key pyrolysis products in real-fuel combustion, and stoichiometric parameters are used to describe their respective amounts. Lumped unimolecular decomposition and H-atom abstraction reactions leading to the aforementioned species are listed as R1 and R2, respectively. The meanings of stoichiometric parameters in R1 and R2 are shown in Table 1: α and β are the numbers of H atoms produced in unimolecular decomposition and H-abstraction reactions for real-fuels, respectively. Furthermore, γ represents the yield of CH₄, λ_3 is the ratio of C₃H₆-to-C₂H₄, $\lambda_{4,1}$ and $\lambda_{4,i}$ are the ratios of 1-C₄H₈-to-C₂H₄ and IC₄H₈-to-C₂H₄, respectively; χ represents the ratio of C₆H₆ to the sum of C₆H₆ and C₇H₈. These stoichiometric parameters have been shown to be weakly dependent on temperature and independent on pressure, and can be obtained by fitting the experimental real-fuel pyrolysis speciation data of CH₄ and C₂H₄ in shock tubes [9,10] and C₃H₆, C₄H₈ isomers, C₆H₆ and C₇H₈ in flow reactors [8,11,12].

R2

where

R represents H, CH₃, OH, O₂, O and HO₂ radicals; $\sum_p = \gamma CH_4 + e_a(C_2H_4 + \lambda_3 C_3H_6 + \lambda_4 C_4H_8) + b_a[\chi C_6H_6 + (1 - \chi)C_7H_8] + \beta H + (1 - \beta)CH_3$

Table 1 Description of independent stoichiometric parameters in the HyChem approach [8]

Parameter	Description
α	Number of H atoms produced in unimolecular decomposition reaction per fuel
β	Number of H atoms produced in H-abstraction reactions per fuel
γ	CH ₄ yield per fuel in addition to H abstraction by CH ₃
λ_3	[C ₃ H ₆]/[C ₂ H ₄]
$\lambda_{4,1}$	[1-C ₄ H ₈]/[C ₂ H ₄]
$\lambda_{4,i}$	[IC ₄ H ₈]/[C ₂ H ₄]
χ	[C ₆ H ₆]/([C ₆ H ₆] + [C ₇ H ₈])

Note: $\lambda_4 = \lambda_{4,1} + \lambda_{4,i}$

Experiments are not always feasible for many research groups and experimental uncertainties should also be fully considered. In addition, the speciation data obtained for one fuel may not be applicable to another fuel; however, the authors of the the HyChem approach have shown that high temperature combustion properties can be predicted with a reasonable degree of accuracy, even when neglecting compositional variability due to the large component number of real-fuels [13]. Nevertheless, this work seeks a more universal approach to account for real-fuel variability without the requirement for speciation data via experimental methods. The present work was inspired by recent progress in predicting ignition propensity based on a chemical functional group approach [14-17], in which the ignition properties were quantitatively correlated to the functional groups. In addition, recent work from this group [18] proved that one or two component surrogates, formulated based on functional groups, can successfully reproduce the combustion characteristics of more complex multicomponent surrogates and real-fuels.

It is hypothesized here that the key pyrolysis products and their stoichiometric parameters in R1 and R2 can be predicted based on the functional groups present in the reacting hydrocarbon fuel. The

goal is to determine whether the quantity and nature of functional groups in a single hydrocarbon molecule or in a complex mixture, are sufficient to establish correlations with the stoichiometric parameters of key pyrolysis products. If this hypothesis is correct, the stoichiometric parameters in a lumped reaction scheme for real fuel pyrolysis can be predicted without experimentation and solely on the basis of the real-fuel's functional group distribution. Furthermore, since the HyChem approach [5] has already shown that high temperature combustion can be reliably predicted once a lumped pyrolysis reaction scheme is developed, it is anticipated that functional groups alone could be enough to predict combustion properties, with an appropriately developed kinetic model.

Several previous studies have systematically investigated the relationship between the fuel molecular structures and intermediate yields. Zamostny *et al.* [19] measured mass fractions of critical pyrolysis products in 56 hydrocarbons at 1083 K, 400 kPa and 0.2-0.4 s, including linear, branched, cyclic, aromatic, unsaturated molecular structures. These speciation data were used to identify important relationships between structural parameters and product yields. According to their studies, the formation of these key pyrolysis products is dependent on the fuels' molecular structures. However, no quantitative structure-yield (i.e. property) relationship was provided by Zamostny *et al.* [19]. Fortunately, the speciation data of key pyrolysis products mentioned above are available in the literature and can also be predicted by kinetic models for various neat fuels, including *n*-alkanes, *iso*-alkanes, cycloalkanes and aromatics, etc., providing the opportunity to develop such quantitative structure-yield relationships, and to predict the stoichiometric parameters for individual components and mixtures.

This work is the first of a two-part study presenting the development of real-fuel kinetic models based on a functional group approach (FGMech) and inspired by the HyChem methodology. In this first study, we address the prediction of stoichiometric parameters for lumped pyrolysis reactions using

functional groups as inputs. In order to achieve the desired goal, we first develop correlations between functional group parameters and stoichiometric parameters for a wide variety of fuel types. To this end, a database consisting of the yield and/or stoichiometric parameters of key pyrolysis products is constructed for various neat fuels, based on previous pyrolysis data and a lumped kinetic model developed by this group in a separate paper [20]. The neat fuels studied include *n*-alkanes, *iso*-alkanes, cycloalkanes and aromatics. Based on the curated datasets, relationships are identified between the stoichiometric parameters and/or yields of key pyrolysis products and the functional groups of their parent fuel. Next, multiple linear regression (MLR) is used to develop quantitative relationships between the stoichiometric parameters and/or yields of key pyrolysis products and the functional groups. The MLR model, trained on neat fuel data, is then validated against experimental speciation data obtained in the pyrolysis of multicomponent surrogate mixtures and real-fuels. The predicted stoichiometric parameters based on the present correlations are compared with those obtained from the HyChem model [8]. Finally, the relationship between the functional groups and stoichiometric parameters is discussed based on the present modeling approach. In part 2 of this study, our findings are extended to develop the FGMeCh kinetic model for predicting real-fuel combustion properties by developing relationships between functional groups and kinetic, thermodynamic, and transport parameters.

2. Data-based approach

A database comprising sets of key products from the pyrolysis of various neat fuels was curated in this work. Figure 1 presents the hierarchical data structure and workflow. Data were curated from two sources: 1) experimental pyrolysis speciation data available in the literature, and 2) stoichiometric parameters of lumped reactions in kinetic models developed in a separate work by this group [20]. The original data for neat fuels with their key pyrolysis products were then divided into four chemical

classes: *n*-alkanes, *iso*-alkanes, cycloalkanes and aromatics, as shown in Fig. 1. Molecular details, such as mass fractions of various functional groups and molecular weights, are included in the *metadata*. After the data were carefully curated and evaluated, they were used for the subsequent development of correlations. Further details on data curation, evaluation, and correlations are discussed in the sections that follow.

Fig. 1 Hierarchical data structure and workflow for the present data-based approach (see text for details).

2.1. Datasets for H, C₂H₄, C₃H₆ and C₄H₈ isomers

The datasets of H, C₂H₄, C₃H₆ and C₄H₈ isomers were curated from a lumped mechanism developed in a separate work by this group [20]. Full details on the construction rules of the lumped mechanism can be found in [20]. USC-Mech II [21] was used as the based mechanism on top of which lumped reactions for neat fuels were added. Briefly, the lumping scheme depends on the steady-state assumption, which requires a time-scale analysis of the underlying reaction processes. H-atom abstraction, isomerization and decomposition reactions are lumped into one reaction step, connecting the fuel molecule directly to its key pyrolysis products. The stoichiometries of key pyrolysis products in the lumped reactions are determined by the equilibrium yields of fuel radicals and the branching ratios of β -C-C bond dissociation reactions of fuel radicals. Over 1000-1500 K and above 1 atm, both

the equilibrium yields and the branching ratios are weakly dependent on temperatures and pressures [5,20,22], so the stoichiometries of these key pyrolysis products are weakly dependent on temperature and pressure. Therefore, these fixed stoichiometric parameters were used for the following correlations.

In [20], the aforementioned approach was followed to develop lumped reaction mechanisms for 50 neat fuels, including C₅-C₁₅ *n*-alkanes, C₅-C₁₀ 2-methyl alkanes, C₆-C₁₁ 3-methyl alkanes, C₈-C₁₂ 4-methyl alkanes, C₇-C₁₁ dimethyl alkanes, C₇-C₁₆ multi-methylated alkanes, cyclohexane and C₇-C₁₁ mono-alkylated cyclohexanes, and C₈-C₁₂ mono-alkylated benzenes. The distribution of datasets from lumped reaction mechanisms for different fuel classes is presented in Fig. 2. The reliability of the lumped mechanisms was validated against experimental data in the literature, including speciation data measured in fuel pyrolysis and oxidation, ignition delay times at high temperatures, and laminar flame speeds [20], indicating that the predicted stoichiometries of H, C₂H₄, C₃H₆, C₄H₈ isomers in the lumped reactions, are reasonably accurate. The stoichiometries of H, C₂H₄, C₃H₆, C₄H₈ isomers in the lumped reactions from our lumped mechanism [20], are listed in the *Supplementary Material*, and were used to develop the correlations that follow.

Fig. 2 Distribution of 50 datasets from lumped mechanism [20] for various fuel classes to obtain stoichiometric parameters of H, C₂H₄, C₃H₆, C₄H₈ isomers.

2.2. Datasets for CH₄, C₆H₆ and C₇H₈

The lumped reactions developed for various fuels in our separate work [20] can only provide the stoichiometric parameters for C₂H₄, C₃H₆, C₄H₈ isomers, as well as H and CH₃. The model cannot provide parameters for CH₄, C₆H₆ and C₇H₈, since these three products are not directly generated via fuel H-atom abstraction reaction and subsequent isomerization and decomposition reactions of fuel radicals. Therefore, we had to figure out another solution to obtaining γ and χ . Our idea was inspired by analyzing the pyrolysis experimental data of Jet A2 reported in [5]. The yield of CH₄ at 1196 K and 12.5 atm in the pyrolysis of Jet A2 almost achieve equilibrium after 1 ms. The normalized peak mole fraction of CH₄ was calculated using Eq. 1, where $x(INT)$ and $x_0(Fuel)$ represent the peak mole fraction of the key pyrolysis product (i.e., CH₄) and the initial mole fraction of the fuel, respectively. Under this condition, the normalized peak mole fraction of CH₄ is around 0.43, which is very close to γ (i.e. 0.45). Furthermore, the yields of C₆H₆ and C₇H₈ also reach equilibrium gradually after 20 ms in the pyrolysis of Jet A2 at 1030 K and 1 atm in a flow reactor. The ratio between the measured peak mole fraction of C₆H₆ and the sum of C₆H₆ and C₇H₈ is 0.51, which is almost the same as χ (i.e. 0.5). Therefore, in this work, we attempted to use the yields of CH₄, C₆H₆ and C₇H₈ in the pyrolysis experiments to obtain γ and χ . It is noticed that γ and χ are independent on temperature and pressure in the HyChem approach [5], while the yields of CH₄, C₆H₆ and C₇H₈ are not. Therefore, the yield of CH₄, and the ratio of [C₆H₆]/([C₆H₆]+[C₇H₈]) obtained in this work are simply used to mimic the values of γ and χ , respectively.

$$x_{normalized} = \frac{x(INT)}{x_0(Fuel)} \quad \text{Eq. 1}$$

The present work first investigated the temperature, pressure and reaction time dependencies of the yields of CH₄, C₆H₆ and C₇H₈, based on detailed kinetic models [23-25]. Three C₁₀ fuels, *n*-decane, decalin and *n*-butylbenzene, were selected to represent paraffinic, naphthenic and aromatic fuel classes.

Four typical experimental conditions were selected from the present curated pyrolysis data for simulation, since 70% of datasets were conducted under these conditions, i.e. (1) $T = 1500\text{ K}$, $P = 0.04\text{ atm}$, $t = 4\text{ ms}$; (2) $T = 1400\text{ K}$, $P = 0.2\text{ atm}$, $t = 20\text{ ms}$; (3) $T = 1300\text{ K}$, $P = 1\text{ atm}$, $t = 110\text{ ms}$; (4) $T = 1083\text{ K}$, $P = 4\text{ atm}$, $t = 300\text{ ms}$. The predicted yields of CH_4 , C_6H_6 and C_7H_8 in the pyrolysis of *n*-decane, decalin and *n*-butylbenzene, under the above four conditions, are shown in Figs. S1-S3 in the *Supplementary Material*. Under all conditions, the yields of CH_4 , C_6H_6 and C_7H_8 are close to achieving equilibrium at the end time. In addition, under the four test conditions, their yields are different, generally within a factor of 2.5. Considering these observed dependencies on temperature, pressure and reaction time, the normalized yields of CH_4 , C_6H_6 and C_7H_8 were calculated under different experimental conditions using Eq. 1, and the mean value was used for the subsequent development of correlations.

For the curation and evaluation of the experimental data, 136 datasets for CH_4 , C_6H_6 and C_7H_8 were initially collected from experiments on the pyrolysis of 44 neat fuels, covering wide temperature, pressure and reaction time conditions. They were obtained in different labs with various experimental apparatus, such as shock tubes [26-29], flow reactors [22,25,30,31] and jet-stirred reactors (JSR) [32,33]. All of these collected speciation data included experimental uncertainties, so they were evaluated to establish consistency across different datasets. In many cases, speciation data of CH_4 , C_6H_6 and C_7H_8 from the pyrolysis of a certain fuel could be obtained from multiple sources. For example, seven datasets were collected for CH_4 from *iso*-octane pyrolysis. As seen from Fig. 3a, the normalized peak mole fractions of CH_4 by the initial mole fractions of *iso*-octane from different sources were compared, and it is apparent that the data from sources NO. 4 and NO. 5 are not in agreement with the others. It was found that these two datasets were collected from the same literature sources [34] and measured with the same apparatus. The conditions investigated are close to those of

the other datasets, but due to these discrepancies compared with other datasets, they were excluded. The mean value of the normalized peak mole fractions from the remaining five datasets was then used for the subsequent correlations.

A consistency check among different intermediates was also performed in this work to remove any anomalies. The consistency check was performed by investigating the relationships between kinetically coupled intermediates. Hansen *et al.* [35] recently performed a data-driven research study, which found that intermediates that are kinetically linked to each other show a linear relationship between their peak mole fractions regardless of the experimental conditions. In the primary decomposition of alkanes, CH₄ concentration is closely related to the yield of CH₃, which accompanies the formation of C₂H₄ from fuel radical decomposition reactions. Therefore, CH₄ and C₂H₄ are kinetically coupled during the primary decomposition of alkanes, and their peak mole fractions should have linear relationship. The quantitative criterion for classifying data as acceptable or anomalous was based on R² of the linear fit. If R² was larger than 0.85, we considered the linear relationship as good, and all the datasets were adopted for developing correlations.

Figure 3b shows the peak mole fractions of CH₄ against those of C₂H₄ for all *n*-alkanes. It was found that most of the datasets obeyed the linear relationship, while one dataset (highlighted in red) is far beyond the trendline. The R² of the linear fit is 0.75, which is lower than the quantitative criterion. This dataset is from *n*-dodecane pyrolysis in a JSR at a residence time of 4.74 s [33]. This long residence time could promote the contributions of secondary reactions, and because of this inconsistency, this dataset was removed from the database. When rejecting the red data point, the R² of the linear fit in Fig. 3b can achieve 0.87 for the remaining datasets, which satisfies the quantitative criterion. Following such evaluation and selection, the final experimental database consisting of 123 datasets from the pyrolysis experiments of 44 neat fuels, was used for the development of correlations.

These final datasets for CH₄, C₆H₆ and C₇H₈ were divided into four chemical classes for the parent fuels. The carbon number range for each class, and the general pressure, temperature and reaction time ranges are briefly summarized in Table 2. Detailed information for each dataset, such as the experimental conditions, uncertainties, reference sources, normalized mole fraction data, etc. can be found in the *Supplementary Material*.

Fig. 3 (a) Normalized peak mole fractions of CH₄ in pyrolysis of *iso*-octane from various literature sources [22,34,36-40]; (b) peak mole fractions of CH₄ as a function of peak mole fractions of C₂H₄ from a variety of pyrolysis datasets of *n*-alkanes. Solid red line represents a linear fit ($y = ax+b$) of the experimental data.

Chemical class	Carbon number	P (atm)	T (K)	Time (ms)	Number of datasets
<i>n</i> -alkane	C ₄ -C ₁₆	1-50	1000-1600	1-300	33
<i>iso</i> -alkane	C ₄ -C ₈	1-60	1023-1500	1-300	23
cycloalkane	C ₆ -C ₁₂	1-200	1075-1587	2-300	38
aromatics	C ₇ -C ₁₀	0.04-50	1083-1700	1-300	29

2.3. Data processing and correlation development

The normalized peak mole fractions of CH₄, C₆H₆ and C₇H₈ were obtained using Eq. 1. For some neat fuels like *n*-heptane and *iso*-octane, there are multiple sources for the speciation data of CH₄, C₆H₆ and C₇H₈. In such cases, the mean normalized peak mole fraction was used in the development of correlations. The stoichiometric parameters for H, C₂H₄, C₃H₆ and C₄H₈ isomers obtained from the

lumped reaction schemes for neat fuels [20] were used in the correlations. Note that the stoichiometric parameters for C₂H₄, C₃H₆ and C₄H₈ isomers are different in unimolecular decomposition and H-atom abstraction reactions [20,22]. In this work, the stoichiometric parameters in H-atom abstraction reactions were used for correlations because of their dominant role in determining the product distributions.

Multiple linear regression (MLR) is a simple, easy to implement statistical technique that can be used to develop an empirical relationship between a dependent variable (y) and multiple independent variables (x_i). In this work, MLR analysis was used to develop the mathematical relationship between the normalized peak mole fractions and/or stoichiometric parameters of the key pyrolysis products (y) and their functional group features (x_i), i.e. the weight percentage of paraffinic CH₃, CH₂, CH, C, naphthenic CH₂, CH, aromatic CH, C, as well as the fuel molecular weight and branching index (BI). BI is a parameter to quantify the effect of branching in *iso*-paraffinic and ringed structures; details can be found in previous work by this group [15]. As seen from Eq. 2, INT represents one of the key pyrolysis product, for example C₂H₄, such that the parameters b₀-b₁₀ can be obtained from MLR by fitting the relationship between 50 stoichiometric parameters of C₂H₄ from 50 neat fuels and the corresponding functional group information of these 50 neat fuels. Given the composition of the functional groups, the average molecular weight and the BI of a fuel mixture, the stoichiometric parameters and/or yields of the key pyrolysis products can be predicted by Eq. 2.

$$\begin{aligned} \text{stoichiometric parameter of INT} = & b_0 + b_1 \times \text{paraffinic CH}_3(\text{wt}\%) + b_2 \times \\ & \text{paraffinic CH}_2(\text{wt}\%) + b_3 \times \text{paraffinic CH}(\text{wt}\%) + b_4 \times \text{naphthenic CH}_2(\text{wt}\%) + b_5 \times \\ & \text{naphthenic CH}(\text{wt}\%) + b_6 \times \text{aromatic CH}(\text{wt}\%) + b_7 \times \text{aromatic C}(\text{wt}\%) + b_8 \times \\ & \text{paraffinic C}(\text{wt}\%) + b_9 \times \text{molecular weight} + b_{10} \times \text{BI} \end{aligned} \quad \text{Eq. 2}$$

3. Results and discussions

3.1. Effect of the functional group on stoichiometric parameter and/or yield of key pyrolysis products

The effect of functional groups on the stoichiometric parameter and/or yield of key pyrolysis products was investigated. C_2H_4 , C_3H_6 and $1-C_4H_8$ are the most abundant primary products in *n*-alkane pyrolysis. Figures 4(a, b) present the stoichiometric parameters of C_2H_4 , C_3H_6 and $1-C_4H_8$ from lumped sub-mechanisms of *n*-alkanes [20] as a function of the weight percent of paraffinic CH_3 and CH_2 , respectively. Results show that the stoichiometry of C_2H_4 is the greatest, followed by C_3H_6 and $1-C_4H_8$ for C_5 - C_{15} *n*-alkanes, consistent with the order of their yields measured by Zamostny *et al.* [19]. With the increasing weight percent of paraffinic CH_3 , the stoichiometry of C_2H_4 , C_3H_6 and $1-C_4H_8$ decrease while with the increasing weight percent of paraffinic CH_2 , the opposite trend is observed, revealing that the C_2H_4 , C_3H_6 and $1-C_4H_8$ stoichiometric parameters increase in the lumped sub-mechanisms of C_5 - C_{15} *n*-alkanes with increasing carbon-chain length. According to the construction principles of the lumped mechanism [20], fuel molecules were all converted to C_0 - C_4 intermediates. Alkenes larger than C_4 were converted to small fragments via H-addition-elimination reactions. Fuel molecules with longer carbon chain lengths have higher fluxes to C_2H_4 , C_3H_6 , $1-C_4H_8$. Therefore, based on the lumped mechanism [20], the stoichiometry of C_2H_4 , C_3H_6 , $1-C_4H_8$ increases with the increasing carbon chain length. However, under practical pyrolysis conditions, larger alkenes ($> C_5$) produced in long chain *n*-alkanes may not be totally converted to small intermediates. Longer carbon chain structures in fuel molecules promote yields of larger alkenes, resulting in the reduction of the carbon flux to C_2H_4 , C_3H_6 and $1-C_4H_8$. Therefore, in practical pyrolysis, their yields may have a turning point, first increasing and then decreasing, as observed in Zamostny's experiments [19].

Fig. 4 Key pyrolysis product stoichiometry/yield as a function of (a) molecular weight percent of paraffinic CH₃ and (b) paraffinic CH₂ in *n*-alkanes, (c) molecular weight percent of paraffinic CH and (d) branching index in *iso*-alkanes, (e) molecular weight percent of naphthenic CH₂ and (f) naphthenic CH in cycloalkanes. 2-Me, 3-Me and 4-Me shown in Fig. 4c represent 2-methylalkane, 3-methylalkane and 4-methylalkane, respectively. 3-MeC₈, 4-MeC₈, di-MeC₈ and tri-MeC₈ in Fig. 4d represent 3-methylheptane, 4-methylheptane, 2,5-dimethylhexane and 2,2,4-trimethylpentane, respectively.

Figures. 4c and 4d present the stoichiometric parameters of C₂H₄ and IC₄H₈ in the lumped sub-mechanism [20] of mono-methylated-alkanes (i.e. 2-methyl-, 3-methyl- and 4-methyl-alkanes) and four C₈ *iso*-alkanes (i.e. 3-methylheptane, 4-methylheptane, 2,5-dimethylhexane and 2,2,4-

trimethylpentane), respectively. Paraffinic CH and BI are unique functional group identities in methylated-alkanes, while they are not present in *n*-alkanes. Increasing CH weight percentage and BI denotes an increasing propensity of substitution in branched alkanes. As mentioned previously, C₂H₄ is the most abundant primary product in *n*-alkanes pyrolysis. As shown in Fig. 4c, increasing weight percent of CH decreases the C₂H₄ stoichiometric parameter. IC₄H₈ is a specific product in *iso*-alkanes pyrolysis. Compared to 3-methyl- and 4-methyl-alkanes, 2-methyl alkanes produce more IC₄H₈ via the H-abstraction reaction on the tertiary carbon-site, followed by a β -scission reaction. Therefore, among the three methylated alkanes, the stoichiometry of IC₄H₈ is largest in the pyrolysis of 2-methyl alkanes. In addition, as shown in Fig. 4d, the stoichiometry of IC₄H₈ among four C₈ *iso*-alkanes increases with the increasing BI, while that of C₂H₄ decreases. BI represents the number and the position of methyl-substitution. Minor changes of the stoichiometric parameters of C₂H₄ and IC₄H₈ are observed between two mono-methylated heptanes while significant differences are present among mono-, di- and tri-methylated C₈ alkanes, indicating the number of methyl-substitution plays a more important role in determining the stoichiometric parameters of C₂H₄ and IC₄H₈, compared with the position of methyl-substitution.

Figures. 4(e, f) present the yields of benzene and toluene as a function of the weight percent of naphthenic CH₂ and naphthenic CH. As noted, these datasets are curated from pyrolysis data in literature. Naphthenic CH₂ and CH are functional groups specific to cycloalkanes. With the increasing weight percent of naphthenic CH₂ and CH, yields of benzene and toluene become less, as seen from Figs. 4(e, f). This trend reveals that longer side-chain structures promote the formation of benzene and toluene. Benzene and alkylated benzenes in cycloalkanes form via two main sources: from the stepwise dehydrogenation reaction, and from the combination of small radicals like C₃H₃ + C₃H₃ [41]. For cyclohexane and C₇-C₁₀ mono-alkylated cyclohexanes, stepwise dehydrogenation reactions make

similar contributions to the production of benzene and alkylated benzenes, since all the fuel molecules have one cyclic ring. However, longer side-chain structures promote the formation of small radicals, resulting in increased yields of benzene and alkylated benzenes in cycloalkanes. Eight aromatic fuels were considered for the correlations of C_6H_6 and C_7H_8 , i.e. toluene, ethylbenzene, *n*-propylbenzene, *n*-butylbenzene, *p*-xylene, *s*-butylbenzene, *iso*-butylbenzene, tetrahydronaphthalene. Since their molecular structures have large differences, monotonous trends of C_6H_6 and C_7H_8 in aromatic fuels are not observed. With the increasing aromatic CH and C, the yields of C_6H_6 and C_7H_8 first increase and then decrease.

3.2. Quantification of functional group effects on the stoichiometric parameter and/or yield of key pyrolysis products

To quantify the effects of the functional group on the stoichiometric parameter and/or yield of key pyrolysis products, an MLR model was developed; correlation results are presented in Table 3, including the size of datasets, regression coefficient (R^2), correlation coefficients (b_0 - b_{10}), and predicted uncertainties. Comparison between the actual and predicted yield of H, C_2H_4 , C_3H_6 , C_4H_8 isomers, C_6H_6 and C_7H_8 can be found in Figs. S4-S6 in the *Supplementary Material*.

Table 3 Correlation results for key pyrolysis product stoichiometry/yield

	[H] ^a _{uni}	[H] ^b _{H-abs}	[CH ₄]	[C ₂ H ₄]	[C ₃ H ₆]	[1-C ₄ H ₈]	[IC ₄ H ₈]	[C ₆ H ₆]	[C ₇ H ₈]
b₀	0.32158	1.19766	0.27845	-0.91518	1.39365	-0.22422	-1.78859	-0.20826	0.26435
b₁	0.16608	-1.26649	-0.71835	-0.95961	-2.36642	0.03716	3.26496	0	-0.47620
b₂	2.03428	0.08207	1.26252	4.33056	-1.15302	0.86594	-0.81829	0.07225	-0.04880
b₃	3.31390	1.88387	4.68755	3.06428	4.22900	0.21365	-2.24257	0.27251	0.36445
b₄	0.97217	0.23891	0.36759	1.91594	-1.19152	0.21220	0.78236	0.12633	-0.21592
b₅	5.06916	-1.82893	4.28712	3.19933	1.03908	-0.10430	-3.55796	0.48545	0.45754
b₆	0	0	0	0	0	0	0	0.31798	0
b₇	0	0	4.27554	0	0	0	0	0.06462	0.37174
b₈	6.22197	4.72354	8.18132	4.87053	6.13616	0.57971	-1.04178	0.58256	1.07356
b₉	-0.00722	-0.00532	-0.00350	0.00040	0.00714	-0.00010	0.01160	0.00161	-0.00061
b₁₀	0.624861	0.02051	-0.33255	0.32640	-1.60319	0.04118	0.06307	-0.04060	-0.01558
datasets size	50	50	40	50	50	50	50	30	27
R²	0.40	0.42	0.82	0.92	0.91	0.87	0.82	0.91	0.86

UF ^c	2	2	2	2	2	2	2	2	2
-----------------	---	---	---	---	---	---	---	---	---

Note:

a and b represent the stoichiometry of H atom from unimolecular decomposition and H-abstraction reaction, respectively.

c: The predicted uncertainty of each dataset was calculated using $\text{Uncert.} = \frac{\text{Predicted value}}{\text{actual value}}$, which was found to be between 0.5 and 2. Therefore, the predicted uncertainty factor (UF) of the actual value [y], means that the upper limit is [y]×UF and the lower limit is [y]/UF.

For each key pyrolysis product, the coefficients b_1 - b_{10} quantify the contributions of corresponding functional group types listed in Eq. 2. For example, b_1 represents the contribution of paraffinic CH_3 . As seen from Table 3, paraffinic and naphthenic functional groups contribute to the stoichiometric parameters and/or yields of all the key pyrolysis products, while aromatic functional groups only contribute to the formation of CH_4 , C_6H_6 and C_7H_8 . Stoichiometries of H atoms, C_2H_4 , C_3H_6 , C_4H_8 isomers are independent of the aromatic functional groups ($b_6 = 0$ and $b_7 = 0$). In the primary decomposition of fuels under pyrolysis conditions, aromatic C-C and C-H bonds are difficult to break, resulting in little contribution from the aromatic functional groups to the yields of these small intermediates. However, for the remainder of key pyrolysis products, the aromatic functional groups can make important contributions. For yields of CH_4 and C_7H_8 , aromatic C (other than aromatic CH ($b_6 = 0$ and $b_7 \neq 0$)), plays an important role, suggesting that the structure of alkylated benzenes contributes significantly to the formation of CH_4 and C_7H_8 . CH_4 and C_7H_8 are mainly from the subsequent reactions of CH_3 and benzyl radicals, respectively. In alkylbenzene pyrolysis, reactions mainly occur on the alkylated side-chain. CH_3 and benzyl radicals can be produced via the unimolecular decomposition reactions and β -scission reactions of fuel radicals on the side-chain. Therefore, aromatic C (other than aromatic CH), makes important contributions to the yields of CH_4 and C_7H_8 . In contrast, for C_6H_6 , both aromatic C and aromatic CH contribute to its yield ($b_6 \neq 0$ and $b_7 \neq 0$). This is because C_6H_6 can not only originate from the ipso-reactions of alkylated benzene relevant to aromatic C, but it can also come from the subsequent reactions of phenyl radical relevant to aromatic CH.

The data curation for CH₄, C₆H₆ and C₇H₈ were from literature pyrolysis experiments. Speciation data for various fuels are limited, especially for C₆H₆ and C₇H₈, which are rarely reported in *n*-alkane and *iso*-alkane pyrolysis. In addition, we followed a rather simple approach of using mean value of yields under various conditions for the correlation development, since more advanced parametric studies require sufficient datasets, which can be the focus of future work. The linear relationships for stable intermediates are generally good ($R^2 > 0.8$), while those for radicals are poor ($R^2 \approx 0.4$), as seen in Table 3. It is found that the yield of H atom fluctuates significantly. Fortunately, the predicted uncertainties of H atoms in both unimolecular and H-abstraction reactions are within a factor of 2, which are close to those of stable intermediates, as seen in Table 3. More advanced non-linear correlation methods, such as artificial neural networks, together with more abundant datasets may be attempted in future work to improve the correlation results.

3.3. Validations of the MLR model against previous pyrolysis data and stoichiometric parameters

To validate the MLR model, speciation data of CH₄, C₆H₆ and C₇H₈ in the pyrolysis of surrogate fuels and real-fuels were collected from the literature [8,42-44]. Shao *et al.* [42,44] measured the speciation data in the pyrolysis of multiple gasoline surrogate fuels comprising *n*-heptane, *iso*-octane, and toluene in a JSR, including CH₄, C₆H₆ and C₇H₈. Figure. 5 presents the MLR model performance for the yield of CH₄ and C₆H₆, in which bars are the experimental data, with well-defined uncertainties, and symbols are the MLR predicted results. The predicted trend of CH₄ yields is in good agreement with the experimental data, as shown in Fig. 5a, while the predicted yields of CH₄ are around 15% lower than the measured results. For the predictions of C₆H₆, the present MLR model predicts its yield well for the pyrolysis of three-component fuels, i.e. TPRF91 (54% C₇H₈/29% *iso*-octane/17% *n*-heptane in mol) and TPRF97.5 (77.5% C₇H₈/8% *iso*-octane/14.5% *n*-heptane in mol), while failing to

capture the trend of C_6H_6 yield in the pyrolysis of binary fuel mixtures, as shown in Fig. 5b. According to the experimental results, the yield of C_6H_6 first increases and then decreases with the decreasing composition of toluene. However, the predicted yield of C_6H_6 monotonically decreases, which is the same as the predicted results of the detailed model developed by Shao *et al.* [44]. The predicted C_6H_6 yields in the pyrolysis of 45% nC_7H_{16} /55% C_7H_8 and 73% nC_7H_{16} /27% C_7H_8 are beyond the experimental uncertainties.

Fig. 5 Measured (bars) [42,44] and predicted (symbols) yields of (a) CH_4 and (b) C_6H_6 in pyrolysis of five gasoline surrogate fuels.

The present MLR model was also validated against pyrolysis experiments for real-fuels. Recently, Pinkowski *et al.* [43] measured speciation data for the pyrolysis of three jet fuels in a shock tube, i.e., jet POSF 10264, POSF 10325 and POSF 10289. The functional group characterization of these three jet fuels was investigated by Lovestead *et al.* [45], using both GC/MS and NMR methods, providing the information needed to quantify different functional groups. Therefore, we can also predict the yields of CH_4 , C_6H_6 and C_7H_8 in real-fuel pyrolysis based on the present MLR model. As seen from Fig. 6, the predicted intermediate yields are in good agreement with those measured in the literature [43].

Fig. 6 Measured (bars) [43] and predicted (symbols) yields of (a) CH₄ and (b) C₆H₆+C₇H₈ in the pyrolysis of Jet POSF 10264, Jet POSF 10325 and Jet POSF 10289.

Following our approach, the stoichiometric parameters defined in the HyChem approach can be easily obtained. Therefore, the stoichiometric parameters of the lumped reactions for three jet fuel pyrolysis were calculated. Figure 7 compares the stoichiometric parameters obtained in the HyChem model [8] using experiments with the present results predicted by the MLR model. The agreement is generally within a factor of two. Note that in Fig. 7b, the predicted $\lambda_{4,i}$ is three times higher than that from the HyChem model, indicating that the stoichiometry of IC₄H₈ is largely over-predicted in this work. The predicted uncertainties in our results could originate from uncertainties associated with quantifying functional groups in real fuels (i.e., reliance of GC/MS and NMR for the structural characterization) and/or the data-based research method.

Fig. 7 Stoichiometric parameters of the lumped reactions for (a) Jet POSF 10264, (b) Jet POSF 10325 and (c) Jet POSF 10289. Detailed descriptions of the stoichiometric parameters can be found in Table 1. Orange and green bars represent the present predicted stoichiometric parameters and those from the HyChem model [8], respectively.

To examine the influence of these stoichiometric parameters on model performance, the HyChem model was updated with the present stoichiometric parameters, while not changing the rate constants and thermodynamic data. Simulations were carried out using Chemkin PRO software [46]. Speciation data measured in shock tubes and flow reactors, as well as the ignition delay times, were simulated with the closed homogeneous batch reactor module. Adiabatic and constant-volume assumptions were selected for simulating shock tube ignition delay times. The maximum production rate of OH* was used to determine the ignition delay times. The species profiles in shock tube experiments were simulated under an adiabatic, constant-pressure assumption, while the speciation in

flow reactor experiments was simulated under a constant-temperature, constant-pressure assumption. Flame speed simulations were performed with the premixed laminar flame-speed calculation module. Thermal diffusion effects were accounted for and mixture-average transport was used.

Figure 8 presents the measured and predicted speciation data in the shock tube and flow reactor pyrolysis of jet POSF 10325 [8]. Dashed lines represent the predicted results from the HyChem model [8], while solid lines are the predicted results of the HyChem model with the updated stoichiometric parameters obtained in this work. The results predicted by the updated HyChem model are similar for most species, however, for IC_4H_8 , as shown in Fig. 8c (purple symbols and lines), the predicted results with the present stoichiometric parameters are much higher than those by the HyChem model. This is because the stoichiometric parameter obtained for IC_4H_8 in this work is higher than that from the HyChem model [8], as shown in Fig. 7b. As for model performance, compared with the experimental data, the HyChem model with the present stoichiometric parameters captures the yield of IC_4H_8 below 15 ms, while over-predicting its yield at longer reaction times. The original HyChem model under-predicts the yield of IC_4H_8 at all reaction times. The model performance with the present stoichiometric parameters was also examined based on global parameters of jet fuels, such as ignition delay times and flame speeds from the literature. As seen from Fig. 9, there is negligible influence in the predictions of laminar flame speeds and ignition delay times.

Fig. 8 Comparison of measured (symbols) [8] and predicted (lines) speciation data in the (oxidative) pyrolysis of jet POSF 10325. (a) Shock tube pyrolysis of 0.73% fuel in Ar at

1228 K and 12.4 atm; (b-c) Oxidative pyrolysis of 314 ppm fuel in a vitiated oxygen-nitrogen mixture in a flow reactor at 1030 K and 1 atm. Solid lines represent the predicted results of the HyChem model with the present updated stoichiometric parameters; dashed lines are those of the original HyChem model [8].

Fig. 9 (a) Comparison between measured (symbols) [8] and predicted (lines) ignition delay times of jet POSF 10325/4% O₂/Ar mixture ($\phi = 1.0$) at $P_5 = 1.3$ and 13.6 atm in a shock tube. (b) Comparison between measured (symbols) [8] and predicted (lines) laminar flame speeds of jet POSF 10325/air mixtures at $T_u = 403$ K and $P_u = 1$ atm. Solid lines represent predicted results of the HyChem model with the present updated stoichiometric parameters; dashed lines are those from the original HyChem model [8].

3.4. Sensitivity analyses on stoichiometric parameters

This section relates the different roles of seven stoichiometric parameters in predicting combustion behaviors. In considering the predicted uncertainties of the stoichiometric parameters by the MLR model (shown in Table 3), model performances were examined by varying each stoichiometric parameter within its uncertainty limits (a factor of two). The results can be found in Figs. S7-S20 in the *Supplementary Material*. Generally, varying each stoichiometric parameter within a factor of two has little influence on the predictions of laminar flame speeds, as seen by the relatively small sensitivity coefficients described below. However, the predictions of ignition delay times and speciation data are more prominently impacted by varying the stoichiometric parameters. Based on

these results, the sensitivity coefficients of speciation data, ignition delay times and laminar flame speeds were calculated to the variation of stoichiometric parameters by using Eq. 3. A and B are the predicted targets (speciation data, flame speeds, etc.) with the increased and decreased stoichiometric parameters, respectively. The sensitivity analysis results are presented in Fig. 10.

$$\sigma_{\text{sensitivity coefficient}} = \log \frac{A}{B} / \log \frac{2.0}{0.5} \quad \text{Eq. 3}$$

Fig. 10 Sensitivity analysis of (a) pyrolysis speciation data in a shock tube; (b) oxidative speciation data in a flow reactor; (c) ignition delay times and (d) laminar flame speeds to seven stoichiometric parameters, i.e. α , β , γ , λ_3 , $\lambda_{4,1}$, $\lambda_{4,i}$ and χ .

As shown in Fig. 10b, the mole fraction of POSF 10325 is sensitive to α and β , which correspond to the formation of H atoms from the fuel's unimolecular decomposition and H-atom abstraction

reactions, respectively. The negative sensitivity coefficients reveal that α and β contribute to fuel consumption. Figures 10(c, d) show that ignition delay times and flame speeds display negative and positive sensitivity to α and β , respectively, indicating that increasing α and β promotes ignition and flame propagation processes. Figures 10(a, b) show that CH_4 , C_3H_6 , 1- C_4H_8 , IC_4H_8 and C_6H_6 mole fractions have high positive sensitivity to γ , λ_3 , $\lambda_{4,1}$, $\lambda_{4,i}$ and χ , respectively, since these parameters correspond to their formation. The formation of toluene (C_7H_8) corresponds to $(1-\chi)$; thus the sensitivity coefficient of C_7H_8 presents a large negative value to χ , as shown in Fig. 10b.

The distributions of small radical species can govern ignition delay times and flame speeds. Larger λ_3 , $\lambda_{4,1}$ and $\lambda_{4,i}$ indicate that more C_3H_6 and C_4H_8 isomers are formed, while less C_2H_4 are generated. Allylic radicals generated in the subsequent oxidation of C_3H_6 and C_4H_8 isomers are less reactive than C_2H_3 or CH_2CHO radicals produced in C_2H_4 oxidation. Therefore, increasing λ_3 , $\lambda_{4,1}$ and $\lambda_{4,i}$ reduces the reactivity of a combustion system. As shown in Figs. 10(c, d), ignition delay times of POSF 10325 display positive sensitivity to λ_3 , $\lambda_{4,1}$ and $\lambda_{4,i}$, while flame propagation of POSF 10325 is negatively sensitive to λ_3 , $\lambda_{4,1}$ and $\lambda_{4,i}$. χ represents the relative amount of benzene and toluene; large χ means more C_6H_6 and less C_7H_8 are produced. Benzyl radical produced in toluene oxidation has a resonance-stabilized structure, which is less reactive. In addition, in flame propagation, the adiabatic temperature also plays an important role in determining flame speeds. C_6H_6 has a higher adiabatic temperature than C_7H_8 , which can promote flame propagation. Therefore, increasing χ promotes the reactivity of the combustion system, resulting in the decrease of ignition delay times and increase of flame speeds, as shown in Figs. 10(c, d).

The sensitivity coefficients of flame speeds are one order of magnitude lower than those of ignition delay times and speciation data. Therefore, the predicted flame speed discrepancies between the present MLR model and the HyChem model are negligible, as shown in Fig. 9b. In contrast,

ignition delay times are sensitive to stoichiometric parameters. The predicted $\lambda_{4,i}$ is higher and β is lower than those from the HyChem model for Jet POSF 10325, which can increase ignition delay time predictions. Furthermore, the lower λ_3 and $\lambda_{4,1}$ and higher α and χ predicted by present MLR model, compared to those from the HyChem model, can decrease predicted ignition delay times. Therefore, the relatively minor differences of the predicted ignition delay times between the two models (see Fig. 9a) are attributed to the compounded canceling effects of varying the stoichiometric parameters.

3.5. Predictions of stoichiometric parameters for surrogate mixtures and real-fuels

Stoichiometric parameters corresponding to lumped pyrolysis reactions for various surrogate and real-fuels were obtained in this work, based on the present MLR model. The notation and formulation of five different jet surrogates are listed in Table 4. Figure 11a presents the predicted stoichiometric parameters for these five jet fuel surrogates (bars), as well as those for Jet POSF 10264 (symbols). The predicted stoichiometric parameters for jet fuel surrogates are very close, and they are also close to those obtained for a real jet fuel, revealing that the radical and intermediate pools in the pyrolysis and high-temperature oxidation of these fuels may be assumed to be similar. This result justifies the use of these proposed surrogate fuels to accurately predict real jet fuel combustion. It is noted that the predicted $\lambda_{4,i}$ for the MURI2 surrogate fuel is much higher than other fuels. Compared with other surrogate fuels in which the *iso*-alkane components are absent, the MURI2 surrogate fuel contains around 30 mol% *iso*-octane, which largely increases the composition of paraffinic CH₃, as well as the BI. Among the five jet fuel surrogates and Jet POSF 10264, the MURI2 surrogate has the largest composition of paraffinic CH₃ and BI. As shown in Table 3, paraffinic CH₃ makes the largest positive contribution to the yield of IC₄H₈ and BI also makes positive contributions. Therefore, the MURI2 jet fuel surrogate would produce more IC₄H₈ under pyrolysis and high-temperature oxidation conditions compared with the other fuels.

Table 4 Formulation of jet surrogates

Notation	Formulation (in mol)	reference
Dagaut surrogate1	74% <i>n</i> -decane/11% <i>n</i> -propylcyclohexane/15% <i>n</i> -propylbenzene	[47]
Aachen surrogate	77.2% <i>n</i> -decane/22.8% 1,2,4-trimethylbenzene	[48]
MURI2 surrogate	40.41% <i>n</i> -dodecane/29.48% <i>iso</i> -octane/7.28% 1,3,5-trimethylbenzene/22.83% <i>n</i> -propylbenzene	[4]
Dagaut surrogate2	70% <i>n</i> -decane/ 30% <i>n</i> -propylbenzene	[47]
Comandini surrogate	30% <i>n</i> -butylbenzene/30% <i>n</i> -propylcyclohexane/40% <i>n</i> -decane	[49]

Fig. 11 (a) Predicted stoichiometric parameters for five jet fuel surrogates (bars), as well as jet POSF 10264 (symbols); The notation and formulation of jet fuel surrogates can be found in Table 4. (b) Predicted stoichiometric parameters for four jet fuels (bars), i.e. jet POSF 10264, jet POSF 10325, jet POSF 10289 and jet POSF 4658, as well as one rocket fuel (large symbols), i.e. RP POSF 5433 and one synthetic aviation jet fuel (small symbols), i.e. S-8 POSF 4734.

Also, stoichiometric parameters corresponding to lumped pyrolysis reactions for real-fuels can be predicted using the present MLR model. Figure 11b presents the predicted stoichiometric parameters of the lumped reactions for six real-fuels. The characterization of these real-fuels can be found in previous studies [45,50,51]. As shown in Fig. 11b, the predicted stoichiometric parameters

for four jet fuels (bars) are very close, which is consistent with the parameters from the HyChem model [8]. In the HyChem model, the stoichiometric parameters obtained for Jet POSF 10264, 10325 and 10289 are the same. This result can be explained by the similar functional group characterization for these different jet fuels. However, the predicted stoichiometric parameters for a rocket fuel (RP POSF 5433) and a synthetic aviation fuel (S-8 POSF 4734) display large discrepancies compared with those for jet fuels, as seen in Fig. 11b (symbols). These two real-fuels mainly contain paraffinic components with few aromatic components, which explains why the C_6H_6 and C_7H_8 yields have large differences compared with the other jet fuels. Paraffinic components produce C_6H_6 and C_7H_8 via the recombination of smaller intermediates, and C_6H_6 is much easier to form than C_7H_8 . Therefore, the yield of C_7H_8 is usually much lower than C_6H_6 in paraffinic rich fuels [52]. The result is that the ratio of C_6H_6 to the sum of C_6H_6 and C_7H_8 , reflecting χ , is higher in these fuels than in the jet fuels. In addition, the larger composition of alkanes—especially *iso*-alkanes in RP POSF 5433 and S-8 4734—results in higher production of CH_4 (i.e. γ) and C_3H_6 (corresponding to λ_3) compared with other jet fuels, as shown in Fig. 11b, indicating that the radical pools of RP POSF 5433 and S-8 POSF 4734 are quite different from those of other jet fuels. In summary, small differences in the quantities of different functional groups results in negligible discrepancies for the predictions of the stoichiometric parameters. However, differences in the types of functional groups present can result in larger discrepancies in the prediction of stoichiometric parameters.

4. Conclusions

This work proposed an alternative approach to obtaining the stoichiometric parameters corresponding to lumped pyrolysis reactions of surrogates and real-fuels. Our initial hypothesis was that stoichiometric parameters in the HyChem lumped modeling approach could be predicted based on the functional groups present in the reacting hydrocarbon fuel. To test our hypothesis, a

methodology was developed to establish functional group correlations with the yields and/or stoichiometric parameters of key pyrolysis products (i.e. H, C₂H₄, C₃H₆, 1-C₄H₈, IC₄H₈, C₆H₆ and C₇H₈) using a database of pure component pyrolysis speciation and a lumped mechanism, as opposed to fitting against real-fuel experimental data. The effects of functional groups on the stoichiometric parameters and/or yield of these pyrolysis products were first identified, and then quantified by MLR analyses. Uncertainties in the present approach could be minimized in future studies by performing high quality pyrolysis experiments (e.g., shock tube pyrolysis [27,52,53]) for a large number of pure components, and then using these for training more advanced non-linear predictive models, such as artificial neural networks.

The product yields, and the stoichiometric parameters for surrogate mixtures and real-fuels, were predicted by the MLR model and validated against pyrolysis data in the literature and the stoichiometric parameters from the HyChem approach. The performance of the HyChem model was also examined by incorporating the present stoichiometric parameters; it was found that the updated model displays performance comparable to the original. The updated model could reasonably predict the speciation data measured in pyrolysis and oxidation experiments, as well as the ignition delay times and laminar flame speeds.

Based on the MLR model, and given functional group features, stoichiometric parameters were predicted for pyrolysis of multiple surrogate mixtures and real-fuels. For different jet fuel surrogates, the predicted stoichiometric parameters are generally close, and they are also similar to those of real jet fuels. However, for real-fuels with different chemical classes, i.e. jet and rocket fuels, the predicted stoichiometric parameters show greater differences. In part 2 of this study, real-fuel lumped kinetic models are developed on the basis of the functional group approach, and stoichiometric parameters obtained in the present work. With such real-fuel lumped kinetic models, the stoichiometric parameters

can be further validated against data in the literature for a wider range of fuels and combustion conditions.

Acknowledgements

The authors would like to thank Prof. Zhandong Wang, Dr. Long Zhao, Mr. Jiabiao Zou and Yan Zhang for sharing their unpublished pyrolysis data, which were extremely useful in the present data-based research. This work was supported by King Abdullah University of Science and Technology (KAUST) with funds allocated to the Clean Combustion Research Center. We gratefully acknowledge contributions from the KAUST Clean Fuels Consortium (KCFC) and its member companies.

Nomenclature

FGMech	functional groups for mechanism development
MLR	multiple linear regression
R^2	regression coefficient
UF	uncertainty factor
GC/MS	gas chromatography–mass spectrometry
NMR	nuclear magnetic resonance
C_6H_6	benzene
C_7H_8	toluene
1- C_4H_8	1-butene
IC $_4H_8$	<i>iso</i> -butene
INT	key pyrolysis product
TPRF91	54% C_7H_8 /29% <i>iso</i> -octane/17% <i>n</i> -heptane in mol
TPRF97.5	77.5% C_7H_8 /8% <i>iso</i> -octane/14.5% <i>n</i> -heptane in mol
Dagaut surrogate1	74% <i>n</i> -decane/11% <i>n</i> -propylcyclohexane/15% <i>n</i> -propylbenzene in mol
Aachen surrogate	77.2% <i>n</i> -decane/22.8% 1,2,4-trimethylbenzene in mol
MURI2 surrogate	40.41% <i>n</i> -dodecane/29.48% <i>iso</i> -octane/7.28% 1,3,5-trimethylbenzene/22.83% <i>n</i> -propylbenzene in mol
Dagaut surrogate2	70% <i>n</i> -decane/ 30% <i>n</i> -propylbenzene in mol
Comandini surrogate	30% <i>n</i> -butylbenzene/30% <i>n</i> -propylcyclohexane/40% <i>n</i> -decane in mol
POSF	fuel designation, not an acronym

References

- [1] C. Westbrook, M. Mehl, W.J. Pitz, G. Kukkadapu, S.W. Wagnon, K. Zhang, Multi-Fuel Surrogate Chemical Kinetic Mechanisms for Real World Applications, *Phys. Chem. Chem. Phys.* 20 (2018) 10535-11486.
- [2] W.J. Pitz, C.J. Mueller, Recent progress in the development of diesel surrogate fuels, *Prog. Energy Combust. Sci.* 37 (2011) 330-350.
- [3] S.M. Sarathy, A. Farooq, G.T. Kalghatgi, Recent progress in gasoline surrogate fuels, *Prog. Energy Combust. Sci.* 65 (2018) 67-108.
- [4] S. Dooley, S.H. Won, J. Heyne, T.I. Farouk, Y. Ju, F.L. Dryer, K. Kumar, X. Hui, C.-J. Sung, H. Wang, M.A. Oehlschlaeger, V. Iyer, S. Iyer, T.A. Litzinger, R.J. Santoro, T. Malewicki, K. Brezinsky, The experimental evaluation of a methodology for surrogate fuel formulation to emulate gas phase combustion kinetic phenomena, *Combust. Flame* 159 (2012) 1444-1466.
- [5] 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.
- [6] E. Ranzi, A. Frassoldati, S. Granata, T. Faravelli, Wide-range kinetic modeling study of the pyrolysis, partial oxidation, and combustion of heavy n-alkanes, *Ind. Eng. Chem. Res.* 44 (2005) 5170-5183.
- [7] E. Ranzi, T. Faravelli, P. Gaffuri, E. Garavaglia, A. Goldaniga, Primary pyrolysis and oxidation reactions of linear and branched alkanes, *Ind. Eng. Chem. Res.* 36 (1997) 3336-3344.
- [8] R. Xu, K. Wang, S. Banerjee, J. Shao, T. Parise, Y. Zhu, S. Wang, A. Movaghar, D.J. Lee, R. Zhao, X. Han, Y. Gao, T. 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.
- [9] X. Han, M. Liszka, R. Xu, K. Brezinsky, H. Wang, A high pressure shock tube study of pyrolysis of real jet fuel Jet A, *Proc. Combust. Inst.* 37 (2019) 189-196.
- [10] J. Shao, Y. Zhu, S. Wang, D.F. Davidson, R.K. Hanson, A shock tube study of jet fuel pyrolysis and ignition at elevated pressures and temperatures, *Fuel* 226 (2018) 338-344.
- [11] K. Wang, R. Xu, T. Parise, J. Shao, A. Movaghar, D.J. Lee, J.W. Park, Y. Gao, T.F. Lu, F.N. Egolfopoulos, D.F. Davidson, R.K. Hanson, C.T. Bowman, H. Wang, A physics based approach to modeling real-fuel combustion chemistry - IV. HyChem modeling of combustion kinetics of a bio-derived jet fuel and its blends with a conventional Jet A, *Combust. Flame* 198 (2018) 477-489.
- [12] Y.J. Tao, R. Xu, K. Wang, J.K. Shao, S.E. Johnson, A. Movaghar, X. Han, J.W. Park, 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 - III. Reaction kinetic model of JP10, *Combust. Flame* 198 (2018) 466-476.
- [13] R. Xu, H. Wang, Principle of large component number in multicomponent fuel combustion – a Monte Carlo study, *Proc. Combust. Inst.* 37 (2019) 613-620.
- [14] K. Dussan, S.H. Won, A.D. Ure, F.L. Dryer, S. Dooley, Chemical functional group descriptor for ignition propensity of large hydrocarbon liquid fuels, *Proc. Combust. Inst.* 37 (2019) 5083-5093.
- [15] A.G. Abdul Jameel, N. Naser, A.-H. Emwas, S. Dooley, S.M. Sarathy, Predicting fuel ignition quality using ¹H NMR spectroscopy and multiple linear regression, *Energy Fuels* 30 (2016) 9819-9835.
- [16] M. Dahmen, W. Marquardt, A Novel Group Contribution Method for the Prediction of the Derived Cetane Number of Oxygenated Hydrocarbons, *Energy Fuels* 29 (2015) 5781-5801.
- [17] A.G. Abdul Jameel, V. Van Oudenhoven, A.-H. Emwas, S.M. Sarathy, Predicting octane number using nuclear magnetic resonance spectroscopy and artificial neural networks, *Energy Fuels* 32 (2018) 6309-6329.
- [18] A.G. Abdul Jameel, N. Naser, A.-H. Emwas, S.M. Sarathy, Surrogate formulation for diesel and jet fuels using the minimalist functional group (MFG) approach, *Proc. Combust. Inst.* 37 (2019) 4663-4671.
- [19] P. Zamostny, Z. Belohlav, L. Starkbaumova, J. Patera, Experimental study of hydrocarbon structure effects on the

composition of its pyrolysis products, *J. Anal. Appl. Pyrolysis* 87 (2010) 207-216.

- [20] X. Zhang, S.M. Sarathy, A lumped kinetic model for high-temperature pyrolysis and combustion of 50 surrogate fuel components and their mixtures, *Fuel* 286 (2021) 119361.
- [21] 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 $H_2/CO/C_1-C_4$ compounds. http://ignis.usc.edu/USC_Mech_II.htm, (2007).
- [22] K. Wang, C.T. Bowman, H. Wang, Kinetic analysis of distinct product generation in oxidative pyrolysis of four octane isomers, *Proc. Combust. Inst.* 37 (2019) 531-538.
- [23] 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, <http://web.stanford.edu/group/haiwanglab/JetSurF/JetSurF2.0/index.html>, (2010).
- [24] M. Zeng, Y. Li, W. Yuan, Z. Zhou, Y. Wang, L. Zhang, F. Qi, Experimental and kinetic modeling investigation on decalin pyrolysis at low to atmospheric pressures, *Combust. Flame* 167 (2016) 228-237.
- [25] Y. Zhang, C.C. Cao, Y.Y. Li, W.H. Yuan, X.Y. Yang, J.Z. Yang, F. Qi, T.P. Huang, Y.Y. Lee, Pyrolysis of n-butylbenzene at various pressures: Influence of long side-chain structure on alkylbenzene pyrolysis, *Energy Fuels* 31 (2017) 14270-14279.
- [26] M.E. MacDonald, W. Ren, Y. Zhu, D.F. Davidson, R.K. Hanson, Fuel and Ethylene Measurements during n-dodecane, methylcyclohexane, and iso-cetane pyrolysis in shock tubes, *Fuel* 103 (2013) 1060-1068.
- [27] K. Yasunaga, H. Yamada, H. Oshita, K. Hattori, Y. Hidaka, H. Curran, Pyrolysis of n-pentane, n-hexane and n-heptane in a single pulse shock tube, *Combust. Flame* 185 (2017) 335-345.
- [28] M.B. Sajid, T. Javed, A. Farooq, Shock tube/laser absorption measurements of methane, acetylene and ethylene during the pyrolysis of n-pentane and iso-pentane, *Combust. Flame* 164 (2016) 1-9.
- [29] M.K. Liszka, K. Brezinsky, Experimental and comparative modeling study of high temperature and very high pressure methylcyclohexane pyrolysis, *Fuel* 237 (2019) 245-262.
- [30] M. Zeng, W. Yuan, W. Li, Y. Zhang, C. Cao, T. Li, J. Zou, Comprehensive experimental and kinetic modeling study of n-tetradecane combustion, *Energy Fuels* 31 (2017) 12712-12720.
- [31] Z. Wang, L. Ye, W. Yuan, L. Zhang, Y. Wang, Z. Cheng, F. Zhang, F. Qi, Experimental and kinetic modeling study on methylcyclohexane pyrolysis and combustion, *Combust. Flame* 161 (2014) 84-100.
- [32] T. Bissoonauth, Z. Wang, S.Y. Mohamed, J.-y. Wang, B. Chen, A. Rodriguez, O. Frotier, X. Zhang, Y. Zhang, C. Cao, J. Yang, O. Herbinet, F. Battin-Leclerc, S.M. Sarathy, Methylcyclohexane pyrolysis and oxidation in a jet-stirred reactor, *Proc. Combust. Inst.* 37 (2019) 409-417.
- [33] O. Herbinet, P.-M. Marquaire, F. Battin-Leclerc, R. Fournet, Thermal decomposition of n-dodecane: Experiments and kinetic modeling, *J. Anal. Appl. Pyrolysis* 78 (2007) 419-429.
- [34] Q. Wang, X. Tang, J. Yang, Y. Zhai, Y. Zhang, C. Cao, J. Wang, Investigations on pyrolysis of isooctane at low and atmospheric pressures, *Energy Fuels* 33 (2019) 3518-3528.
- [35] N. Hansen, X. He, R. Griggs, K. Moshhammer, Knowledge generation through data research: New validation targets for the refinement of kinetic mechanisms, *Proc. Combust. Inst.* 37 (2019) 743-750.
- [36] W. Shen, Y. Zhang, B. Zhao, D. Chang, J. Lyu, H. Zhang, Experimental and modelling studies on pyrolysis and its main light gas products for typical C8 hydrocarbons, *J. Anal. Appl. Pyrolysis* 142 (2019) 104622.
- [37] K. Yasunaga, S. Etoh, H. Yamada, H. Oshita, Y. Hidaka, Modeling and experimental study on pyrolysis of isooctane and n-heptane behind reflected shock waves, *Chem. Lett.* 47 (2018) 747-750.
- [38] W. Shen, Y. Cao, Y. Zhang, H. Zhang, Experimental and modeling study of chain alkanes pyrolysis, *J. Combust. Sci. Technol.* 24 (2018) 446-450.
- [39] T. Malewicki, A. Comandini, K. Brezinsky, Experimental and modeling study on the pyrolysis and oxidation of isooctane, *Proc. Combust. Inst.* 34 (2013) 353-360.
- [40] D.F. Davidson, M.A. Oehlschlaeger, R.K. Hanson, Methyl concentration time-histories during iso-octane and n-heptane oxidation and pyrolysis, *Proc. Combust. Inst.* 31 (2007) 321-328.

- [41] Z. Wang, Z. Cheng, W. Yuan, J. Cai, L. Zhang, F. Zhang, F. Qi, J. Wang, An experimental and kinetic modeling study of cyclohexane pyrolysis at low pressure, *Combust. Flame* 159 (2012) 2243-2253.
- [42] C. Shao, H. Wang, N. Atef, Z. Wang, B. Chen, M. Almalki, Y. Zhang, C. Cao, J. Yang, S.M. Sarathy, Polycyclic aromatic hydrocarbons in pyrolysis of gasoline surrogates (n-heptane/iso-octane/toluene), *Proc. Combust. Inst.* 37 (2019) 993-1001.
- [43] N.H. Pinkowski, S.J. Cassady, D.F. Davidson, R.K. Hanson, Spectroscopic inference of alkane, alkene, and aromatic formation during high-temperature JP8, JP5, and Jet-A pyrolysis, *Fuel* 269 (2020) 117420.
- [44] C. Shao, G. Kukkadapu, S.W. Wagnon, W.J. Pitz, S.M. Sarathy, PAH formation from jet stirred reactor pyrolysis of gasoline surrogates, *Combust. Flame* 219 (2020) 312-326.
- [45] T.M. Lovestead, J.L. Burger, N. Schneider, T.J. Bruno, Comprehensive assessment of composition and thermochemical variability by high resolution GC/QToF-MS and the advanced distillation-curve method as a basis of comparison for reference fuel development, *Energy Fuels* 30 (2016) 10029-10044.
- [46] ANSYS Chemkin-Pro: a chemical kinetics package for analysis of gas-phase chemical kinetics. Reaction Design, 2019 R2.
- [47] P. Dagaut, A. El Bakali, A. Ristori, The combustion of kerosene: Experimental results and kinetic modelling using 1- to 3-component surrogate model fuels, *Fuel* 85 (2006) 944-956.
- [48] S. Honnet, K. Seshadri, U. Niemann, N. Peters, A surrogate fuel for kerosene, *Proc. Combust. Inst.* 32 (2009) 485-492.
- [49] A. Comandini, T. Dubois, N. Chaumeix, Laminar flame speeds of n-decane, n-butylbenzene, and n-propylcyclohexane mixtures, *Proc. Combust. Inst.* 35 (2015) 671-678.
- [50] T.M. Lovestead, B.C. Windom, J.R. Riggs, C. Nickell, T.J. Bruno, Assessment of the compositional variability of RP-1 and RP-2 with the advanced distillation curve approach, *Energy Fuels* 24 (2010) 5611-5623.
- [51] M.L. Huber, B.L. Smith, L.S. Ott, T.J. Bruno, Surrogate mixture model for the thermophysical properties of synthetic aviation fuel S-8: Explicit application of the advanced distillation curve, *Energy Fuels* 22 (2008) 1104-1114.
- [52] T. Malewicki, K. Brezinsky, Experimental and modeling study on the pyrolysis and oxidation of n-decane and n-dodecane, *Proc. Combust. Inst.* 34 (2013) 361-368.
- [53] R.K. Hanson, D.F. Davidson, Recent advances in laser absorption and shock tube methods for studies of combustion chemistry, *Prog. Energy Combust. Sci.* 44 (2014) 103-114.