energy&fuels

pubs.acs.org/EF

Article

GCxGC-Based *iso*-Alkane Subgrouping for Enhanced Compositional Analysis of Sustainable Aviation Fuels

Published as part of Energy & Fuels special issue "18th International Conference on Stability, Handling and Use of Liquid Fuels (IASH2024)".

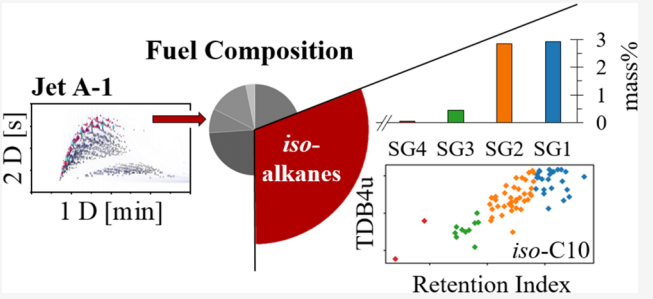
Hannes Lüdtke,* Florian Pütz, Samuel Grams, Uwe Bauder, Markus Köhler, Andreas Huber, and Patrick Oßwald

Cite This: *Energy Fuels* 2025, 39, 4434–4449

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: To evaluate sustainable aviation fuels (SAF) and other novel jet fuels, nontargeted comprehensive analysis by two-dimensional gas chromatography (GCxGC) is commonly utilized. The obtained results are necessary for subsequent model-based prescreening applications. The uncertainty of the respective property predictions is dependent on the degree of compositional detail as some properties (e.g., flash point, freezing point) are strongly influenced by structural molecular features. In the absence of sufficient structural reference data, individual identification from reference databases is usually not possible. Consequently, the results obtained from GCxGC are generally categorized by the carbon number and group type. To obtain greater details on the isomeric structure distribution of fuels, the *iso*-alkane family was further investigated. Therefore, a multilinear regression model for *iso*-alkane retention indices (RI) was constructed from molecular descriptors. Subsequently, the combined database from measurement and literature was extended by prediction to complete the data for all possible 42,900 branched isomers within the jet range (C7-C17). The isomeric structures were sorted into subgroups by their respective retention behavior, thereby correlating with the present number of molecular branches. The structural subgroup information was then used to create branching indicators to quantify and compare the subgroup distributions of different fuels. It was evident that isomeric distributions were unique to the respective samples. The new detail of composition will aid the characterization and differentiation of different fuels and present further potential for fuel assessment.

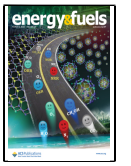


1. INTRODUCTION

Sustainable aviation fuels (SAF), i.e., fuels produced from nonfossil, renewable feedstocks, and renewable energy, are one of the key strategies to reduce the climate impact of aviation. In addition to reducing net carbon emissions, SAF has great potential to significantly reduce aviation's non-CO_2 climate impacts[1−3] by specific compositional fuel design.[4−6] The EU will introduce a mandatory regulation of minimum 2% of SAF from 2025 on, with increasing amounts to 70% already in 2050.[7] To accommodate this and other upcoming regulations, a variety of international projects was and is focused on creating, investigating, and evaluating new production pathways for SAF to ramp up the production capacities.[8−10] Currently, 11 pathways are already approved by ASTM (American Society for Testing and Materials): 3 pathways for coprocessing with conventional crude oil according to ASTM D1655 − 24 and 8 production pathways for synthetic blending components according to ASTM D7566 − 24B. The latter includes, among others, processes for synthetic paraffinic kerosene from hydroprocessed esters and fatty acids (HEFA-SPK), alcohol to jet (ATJ-SPK), or Fischer−Tropsch (FT-SPK).[11]

A significant proportion of the already approved and currently discussed SAF routes, including the aforementioned processes, contain high paraffinic compositions (often mainly branched *iso*-alkanes). Some routes are even discussed to be applied as up to 100% aliphatic non-drop-in fuels in upcoming ASTM specifications since they significantly reduce non-CO_2 effects.[12−14] Therefore, the detailed characterization of the *iso*-alkane fraction in such jet fuel candidates is becoming increasingly important for a refined understanding and prediction of fuel properties[15] and their chemical reaction network.[16−18]

Received: November 14, 2024
Revised: January 31, 2025
Accepted: January 31, 2025
Published: February 25, 2025



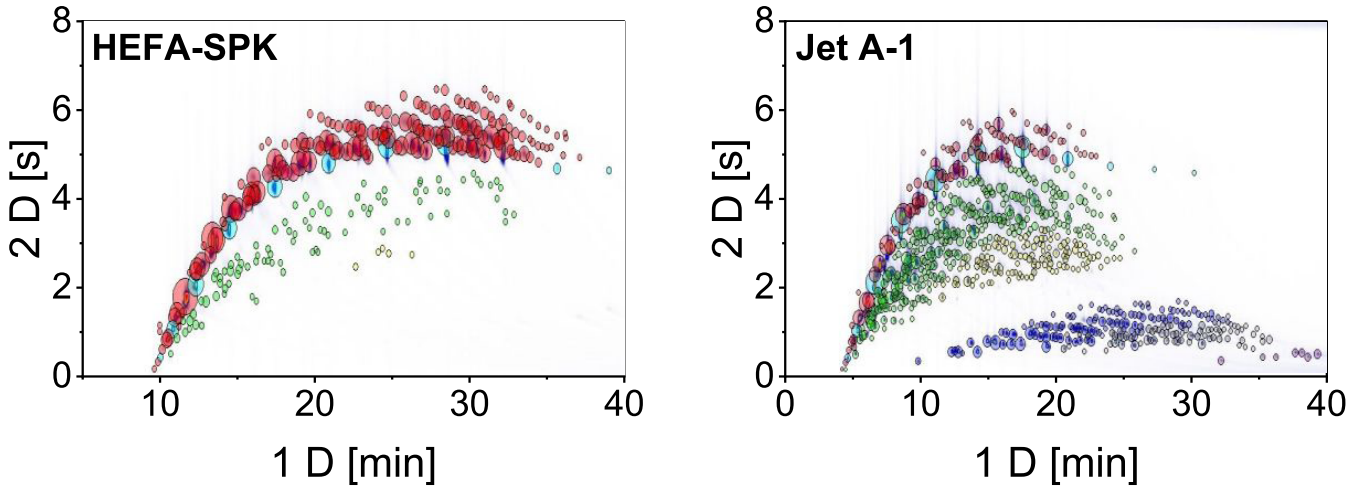


Figure 1. GCxGC chromatograms of HEFA-SPK and Jet A-1 fuel samples. The detected group types are described by color: *n*-alkanes (light blue), *iso*-alkanes (red), *cyclo*-alkanes (green), *dicyclo*-alkanes (yellow), monoaromatics (blue), cycloaromatics (gray), diaromatics (purple).

In order to support manufacturers in the development of new sustainable aviation fuels, the prediction of fuel properties is an important step that is often applied before the ASTM D4054 evaluation process.[19−26] The well-known prescreening process developed by Heyne, Rauch et al. combines several essential property prediction models with a range of experience and evaluation on certification boundaries, to derisk certification.[27−29] This process consists of two tiers, tier α and tier β. Tier α is applied to maximize information from minimal sample volume since the development of novel fuels is typically initiated at the laboratory scale. The detailed chemical composition of the fuel is obtained and utilized to model and predict essential fuel properties, relevant for the safety and performance of standard operation. The compositional information is determined through two-dimensional gas chromatography (GCxGC), which allows high separation efficiency and resolution of several hundred to thousand different fuel components.[30]

Example chromatograms of a HEFA-SPK and a Jet A-1 can be seen in Figure 1. Throughout the development of prescreening, there have been multiple different model approaches used to process the GCxGC results with low uncertainties. In the end, the quality of property predictions remains strongly dependent on the overall detail of composition.[15,25,27,31]

GCxGC allows for a significantly higher resolution than any one-dimensional (1D) technique by adding another separation mechanism with a second column material. The second dimension is mainly required for the separation of different group types (G-T). Additionally, it can improve the separation of compounds from the same G-T but different carbon number (*nC*). Despite the high resolution, single-compound identification remains a challenge. Therefore, common GCxGC data assessment usually only summarizes the separation of signals by *nC* and G-T.[32−34] Hall et al. already showed that some properties are mainly dependent on those two categories (e.g., heat of combustion). However, other properties are significantly influenced by the isomeric structures. In particular, the viscosity and freezing point were reported to be most impacted by the isomeric structure in comparison to molecular size and G-T.[15] Hence, summing up all signals of the same *nC* and G-T inherently introduces some uncertainties for property predictions, since the diversity of isomeric structures increases exponentially with *nC* (Table 1). As a result, the range of certain properties within these isomeric groups can increase as well. This work starts to examine the separation and quantification of structural features within the typical GCxGC resolved hydrocarbon groups and highlights the potential implications for fuel assessment. Given the focus on *iso*-alkane components in the currently approved SAF routes, the proof of concept herein is based on the separation analysis of this hydrocarbon family specifically.

Table 1. Possible Branched *iso*-Alkane Structures in the Jet Fuel Relevant Carbon Range *nC7*-*nC17*[a]

carbon	possible Isomers	MS data	RI data
7	8	8	8
8	17	17	17
9	34	34	34
10	74	39	74
11	158	20	28
12	354	29	42
13	801	45	43
14	1857	13	28
15	4346	10	22
16	10,358	24	32
17	24,893	8	26
jet fuel range	42,900	247	354

[a]Calculated with MOLGEN 5.0.[37] Limited availability of reference data for mass spectrometry fragmentation patterns (MS) and gas chromatography retention indices (RI) is apparent, starting from C10.

2. METHODOLOGY

This study focuses on increasing the compositional detail of comprehensive two-dimensional gas chromatography (GCxGC) for fuel analyses. Retention index (RI) modeling of *iso*-alkanes is carried out to evaluate the differences of retention behavior between isomeric structures. The study highlights the possibility of grouping *iso*-alkanes of similar substructure within the chromatographic separation space of each carbon number. Additionally, the definition of branching indicators allows a quantitative description of isomeric distributions within different fuels. To demonstrate the potential of this methodology for fuel assessment, subgroup

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

and branching indicators are determined for eight fuel samples (Figure 2).

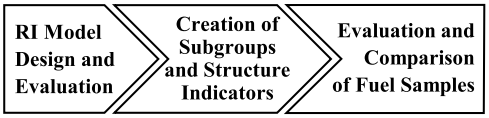


Figure 2. Schematic representation of the main steps in the paper.

2.1. GCxGC—Instrumentation. Measurements were conducted on a two-dimensional gas chromatography system equipped with a flame ionization detector and a quadrupole mass spectrometer (GCxGC-FID/qMS, Figure 3). The system consisted of two ovens (GC2010; GC2010 Plus; Shimadzu) connected by a heated transfer line. The first-dimension column and the second-dimension column were heated independently (see the method parameters in Table 2). This allowed for more efficient use of the separation space, reduction of coelution, and individual consideration of column stabilities and sensitivities. Modulation between both dimensions was performed with a cryogenic loop modulator (ZX1, Zoex) through periodic liquid nitrogen application. This enabled sharp signals, even with the most volatile compounds in typical jet samples. The setup allowed for parallel detection with FID and MS through a four-port splitter (Trajan, SilFlow 4 port splitter). Through auxiliary pressure application of additional carrier gas and guard column dimensions, the flow ratio was configured (calculated by Advanced Flow Technology, Shimadzu). The FID allows for precise quantification with a wide linearity range for different hydrocarbon families. The qMS, equipped with electron impact ionization (EI), allows further structural and group-type differentiation of the present sample components. Especially for novel fuel types, uncommon impurities or byproducts can be more easily identified without universal calibration. Data acquisition was performed by using Shimadzu LabSolutions GCMS for both detectors. Sample analysis and quantification was processed with GCImage software (Zoex, v2022r3). Quantification was performed on the FID signal. Identification of hydrocarbon classes and carbon numbers was carried out for each sample individually by combined evaluation of retention behavior and MS fragmentation data. After integrating signals over their respective areas and assigning them a certain G-T and *nC*, they were quantified by the respective created relative response factors for the FID. Detector response differences and possible drifts in detection ratio were considered through the list of standard substances calibrated over the separation space. Further details about data evaluation and detector calibration were already described in Melder et al.[33]

Table 2. GCxGC-FID/qMS Chromatographic Conditions and Detection Parameters

	GCxGC-FID/qMS
injection	0.3 μL neat sample; split 300:1 at 330 °C
oven ramp 1	55−260 °C; 3.5 °C/min
column 1	HeavyWAX
	60 m × 0.25 mm; film: 0.25 μm
oven ramp 2	40−330 °C; 3.5 °C/min
column 2	BPX1
	3 m × 0.1 mm; film: 0.1 μm
carrier gas flow rate	16 cm/s; constant velocity
modulation parameters	8 s; 0.5 s hot pulse (380 °C)
guard columns	MS: 0.4 m; ID: 0.15 mm
	FID: 0.47 m; ID: 0.2 mm
	4 Port Splitter
auxiliary pressure	48 kPa
flow ratio	MS 1: 1.45 FID
ion source temperature	250 °C
transfer line temperatures	330 °C
acquisition frequency	FID (200 Hz); MS (50 Hz)

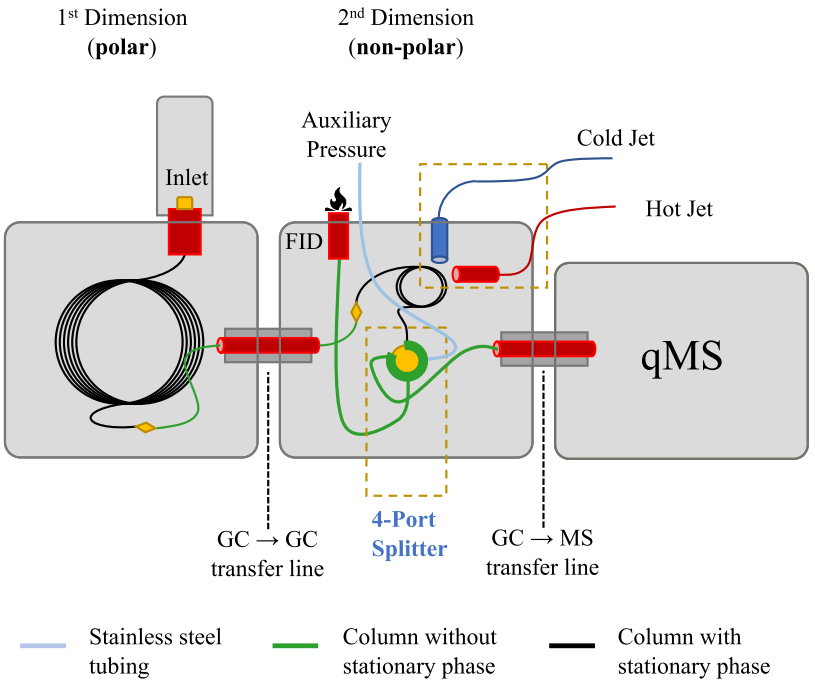


Figure 3. Schematic depiction of dual oven GCxGC-FID/qMS instrumentation.

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

2.2. Retention Index and Retention Behavior. The retention behavior of individual isomeric fuel compounds can be used to identify individual structures by position in the chromatogram. A compounds retention is intercomparable for different systems through the application of the retention index (RI). This index basically contextualizes the relative retention time of a component with respect to a series of reference compounds under the same experimental conditions. The usual reference series consists of *n*-alkane standards that define the coordinate system by $nC \times 100 = \mathrm{RI}$. For linear temperature programming, the RI of a compound z is calculated as described in eq 1.[35]

$$\mathrm{RI}_z = 100 \cdot \left[\frac{t_{Rz} - t_{Rn}}{t_{R(n+1)} - t_{Rn}} + nC \right] \quad (1)$$

t_{Rn} = retention time *n*-alkane
$t_{R(n+1)}$ = retention time n + 1-alkane
t_{Rz} = retention time compound z
nC = carbon number of *n*-alkane

The last two digits indicate the relative positions of the respective species between the corresponding *n*-alkanes. By this logic, a compound with RI = 1160 would be detected 60% of the way between undecane (RI = 1100) and dodecane (RI = 1200).

The two-dimensional chromatography setup is essential for the separation and identification of different group types and different molecular weights. To compare compounds with the same nC and within the same G-T, first-dimension retention is generally sufficient. Therefore, RI data from one-dimensional setups could also be used to investigate structure-to-retention behavior of isomers within one G-T of GCxGC chromatograms. Precise and reliable identification of specific compounds requires reference data. This can either be obtained from actual measurements of known species or the application of (commercial) literature databases (e.g., NIST17[36]). However, the available databases for both MS and RI entries are limited, in contrast to the number of possible branched isomers of the *n*-alkanes. The number of possible isomeric structures is summarized in Table 1, arranged in accordance with the respective chain length for *iso*-alkanes. Isomers outside the designated carbon range ($nC7$-$nC17$) were excluded from detailed structural discussion, as isomeric diversity of $nC18$ (60,522 possible structures, MOLGEN[37]) would already be more than twice the structural isomers of the herein considered jet fuel range. Nevertheless, the expansion of structural analysis outside the model range will be a part of future investigations. Isomeric diversity below $nC7$ only includes 7 individual isomers in between $nC4$-$nC6$ and could therefore be manually distinguished, if necessary. Data for isomeric structures are derived from calculations using the MOLGEN 5.0 software.[37]

The relative retention behavior, i.e., the order of elution from a given column, is expected to be mainly influenced by material of the GC columns and is therefore reproducible in other instruments.[38] For the characterization of the retention behavior within the present chromatographic setup, 44 commercially available *iso*-alkanes were acquired and tested on the system. The limited availability of heavy isomers resulted in an overrepresentation of low-carbon molecules. For the further investigation of RI and structure relationships, the RI database of NIST17[36] (with MS Search 2.3 software) was used. Listed errors from the database were not considered as most literature values were reported without uncertainty, and when available, they were generally below 5 RI (mean error below 2 RI). Despite application of a polar column, the correlation of RI's of measured *iso*-alkanes and polar database values (NIST17[36]) differed significantly (R^2 = 0.9107). Furthermore, the limited availability of polar RI's in the jet range provided no representative database (75 RI's of branched isomers between $nC7-17$) relative to the overall number of possible isomers. Nonpolar (100% polydimethylsiloxane columns, PDMS) literature RI values were more commonly reported and were found to be in significantly closer correlation to the measured isomers (R^2 = 0.9937). Additionally, the correlation of nonpolar (275 RI's) and semistandard nonpolar RI's (95% PDMS; 207 RI's); (R^2 = 0.9997) was significant. This demonstrated that the literature RI values for nonpolar and semi-nonpolar columns were sufficiently representative of the instrumentation used. Consequently, both lists were merged into a combined RI database for the *iso*-alkane components. A total of 354 RI values for individual branched isomers between $nC7-17$ (see distribution in Table 1) were selected from NIST17.[36] For isomeric structures listed in both data sets, the mean value was used in the following procedures.

2.3. Retention Index Model. Since RI and MS databases are still limited, identification of the specific molecular structure in the *iso*-alkane groups is barely possible. To overcome the lack of reference data, a possible approach to this problem is to create an RI model. Different developments have been proven to provide precise RI predictions. Some models partially relied on other experimental properties or the simulation of them (e.g., vapor pressure,[39] solvation parameters[40]). However, this only shifts the limitations to the quality of other models or a database of other properties. Other approaches tried to develop thermodynamic models by extensive temperature-retention calibration and functional-group impact studies within molecules.[41,42] This approach was rejected, as it was expected that the amount of available isomeric structures for setup calibration would not be able to sufficiently describe the possible structural diversities within larger molecules. In order to obtain a complete data set of all possible isomeric structures, a useful approach is the usage of molecular descriptors that are calculatable from the respective structural configuration. This is referred to as quantitative structure−property relationships modeling (QSPR) and allows the necessary values to be gathered universally. Zhang et al. for example reported a promising multiple linear regression (MLR) QSPR approach for hydrocarbon groups in gasoline.[43] Yet, they did focus on group-type-independent effects and overall signal differentiation. Katritzky et al. even underlined the applicability for branched alkanes, although exclusively for methylated compounds and with focus on $nC > 20$.[44] Thus, the MLR approach was investigated with a focus on the identification of intragroup effects of *iso*-alkanes within the typical jet fuel range $nC7-17$.

To find the optimal MLR variables, 1875 molecular descriptors were calculated from SMILES keys (simplified molecular-input line-entry system) of the available data set isomers. This was done with the PaDELPy[45] and RDKit[46] modules for python, starting from already existing parts of the work carried out by Hall et al.[15] Molecular descriptors were tested on correlation with the isomers RI using Pearson,[47] Spearman,[48] and Kendell Tau Tests[49] over the whole data set and separately on the RI subsets of each individual carbon

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

number (nC). The threshold of significant correlation was set to $R^2 \geq 0.8$, to reduce the number of descriptors to be evaluated for the model approach. It was evident that significant correlation over the whole data set was sometimes already achieved by the fact that molecules with higher nC generally elute at later RI's. In total, over 500 descriptors had $R^2 \geq 0.8$ over the whole data set. Yet, a significant proportion of those descriptors did not correlate within the subsets of individual nC and therefore had no structural impact other than the molecular size. Since differentiation by structural features was the main interest of the model, only the gathered significant descriptors ($R^2 \geq 0.8$) throughout the nC-subset correlations were subsequently evaluated.

Despite the fact that the nC subsets were limited by the number of available datapoints, it ensured the focus on the evaluation of intragroup structural effects. With a specific focus on correlation results for *iso*-alkanes of the same size, molecular descriptors were counted for and sorted by frequency of significant correlation within $nC7$-$nC17$. Certain nC subsets were found to correlate with less than 10 descriptors, but it was possible to identify promising descriptors that often reoccurred for different nC-sets. Pearson correlations showed the most explicit results. Specifically, the topological distance-based (TDB) descriptors significantly correlated throughout all RI subsets of different molecular sizes. Among those, TDB4u (three-dimensional (3D) topological distance-based autocorrelation - lag 4/unweighted) was found to be present in 8/11 Pearson subset correlations for $nC7$-$nC17$ and was chosen as the first parameter to be tested for the model approach.

The TDB descriptors, first introduced by Klein et al.,[50] describe the relationship of the average three-dimensional Euclidean distance (r_{ij}) between possible atom pairs in a molecule, in relation to the path length (k, i.e., the number of bonds). It basically compares the 3D compactness of different molecules within a specific range of atoms.

$$\mathrm{TDBku} = \frac{1}{\Delta_k} \cdot \sum_{i=1}^{A-1} \sum_{j=i+1}^{A} (R_i^{\mathrm{cov}} \cdot r_{ij} \cdot R_j^{\mathrm{cov}}) \cdot \delta(d_{ij};\, k) \tag{2}$$

k: path length
k: path length
A: number of atoms
Δ_k: number of atom pairs located at d_{ij}
d_{ij}: topological distance, min. number of connecting bonds between i and j atoms
r_{ij}: geometric distance between ith and jth atoms in the three-dimensional structure
R^{cov}: covalent atom radius of i and j atoms
δ: 0 if $d_{ij} \neq k$; 1 if $d_{ij} = k$

The TDB4u descriptor, as defined by eq 2, exhibited a strong nC-dependent correlation to the RI (Figure 4). To build a generally applicable MLR with the inclusion of the total available data set, the nC was chosen as the second parameter in the model approach.

Further analysis showed that with increasing nC a rising variance in the relation of TDB4u to RI appeared (Figure 4, larger nC can be found in Figure 6). To account for the deviation from functionality, additional structural features were evaluated within those nC groups. Thereby, the number of branches [CX4H3] and the type of carbon centers (quaternary, [CX4H0], i.e., the carbon center is connected to four other carbons and no hydrogen) were found to describe certain effects on the RI (Figure 5). Both descriptors were further considered as third and fourth factors for the model approach.

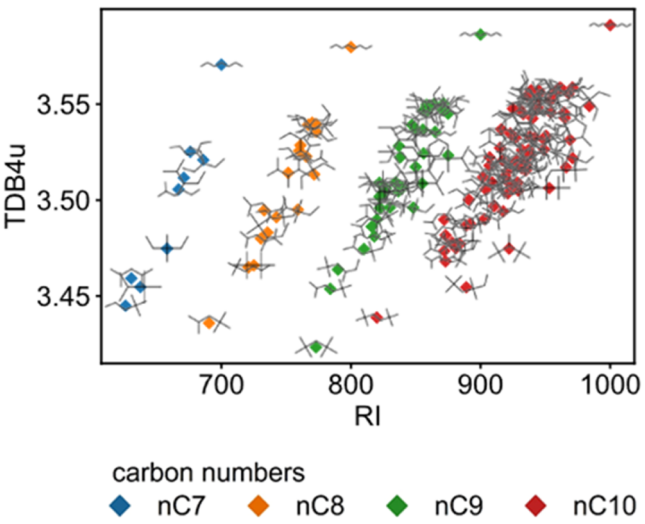


Figure 4. Correlation of TDB4u with the available retention indices from $nC7$-$nC10$. Individual molecular structures are indicated as a small overlay for each data point; the color illustrates the respective nC.

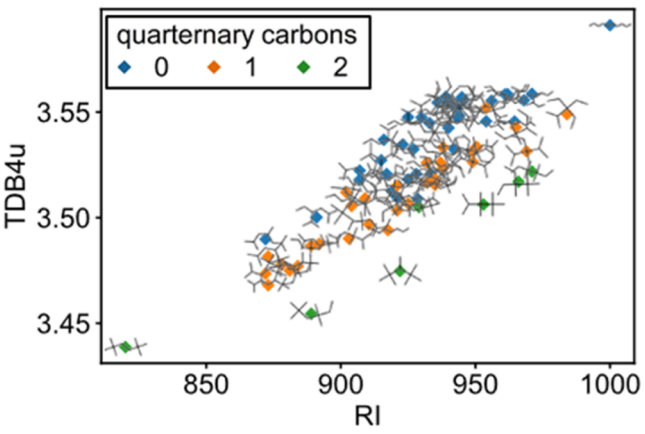


Figure 5. Illustration of intragroup correlation within the TDB4u descriptor to RI relations for $nC10$ isomers. The number of quaternary carbon centers [CX4H0] is indicated by different colors.

The increase of molecular weight and consequently nC, led to an under-representation of long-range effects by the TDB4u descriptor, as only four-step pathways are considered. With higher nC however, longer paths than 4 bonds in between different branches are increasingly likely. Due to the reduced deviation of TDB4u, despite the existing RI differences, it was determined that isomers with high nC were not adequately described with TDB4u (Figure 6). To account for different effects in larger molecules, the potential necessity for a longer path length was investigated. The plotting of TDB7u versus RI showed an improvement in the correlation between TDB and RI among isomers with formerly almost constant TDB4u values.

It was apparent that long-range interactions (described by TDB7u) were not applicable for small molecules, while the small-range interactions (described by TDB4u) created higher uncertainties for larger isomers. To account for the respective limitations of the selected distance for the TDB descriptors,

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

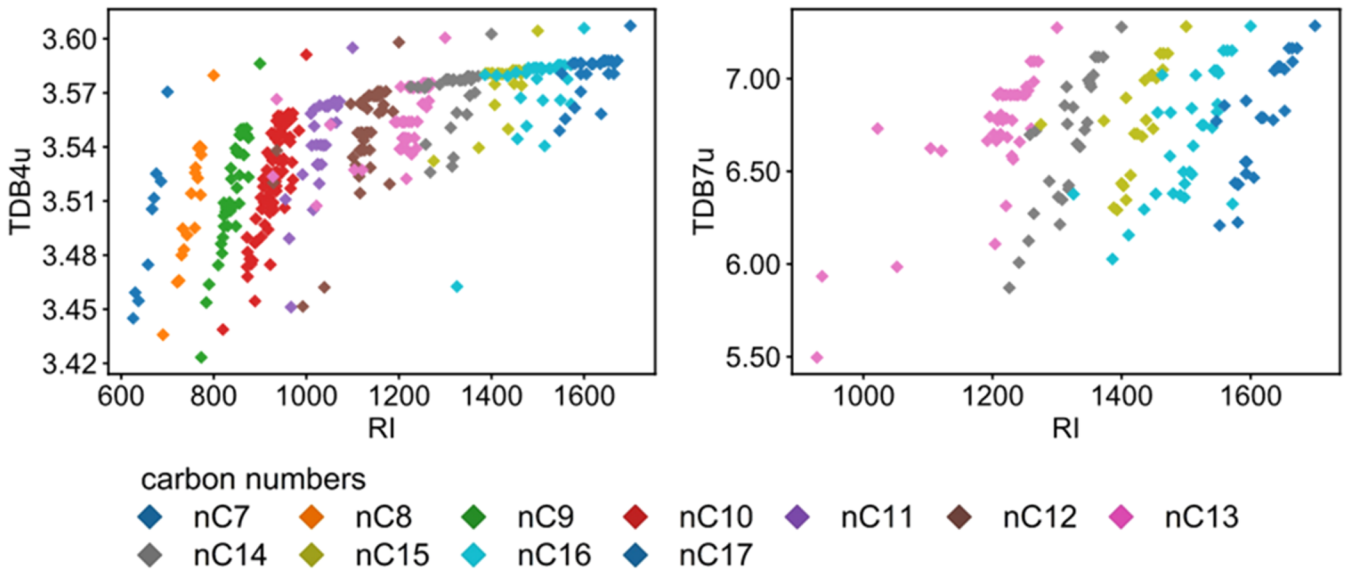


Figure 6. Illustration of carbon number-dependent quality of correlation between TDB4u and RI (left). Change of correlation with higher pathway consideration between TDB7u and RI for high *n*C isomers (right).

the data set was divided into two models based on molecular size. The low carbon numbers (lowC, *n*C7-*n*C12) were fitted with TDB4u. The higher carbon numbers (highC, *n*C13-*n*C17) were modeled with a combination of TDB6u and TDB7u. Both MLR approaches further included the carbon number *n*C, the number of methyl groups [CX4H3], and the number of quaternary carbon centers [CX4H0]. No adjustments beyond the division into two models were considered, to avoid further reduction of the training data set sizes (lowC 203 RI; highC 151 RI).

Due to the limited availability of datapoints, both models were also fitted using the full respective data sets to generate coefficients. The Ridge Regression Cross Validation model from Python's SciKit Learn[51,52] library was applied to fit the data. Ridge regression was chosen to account for collinearity and prevent overestimation of the impact for single parameters. This is done by incorporation of a regularization term, which reduces the variance of the estimation by penalizing large coefficients.[53] Only for estimation of the precision of the model's predictions, the data sets were randomly divided into training (80%) and test sets (20%). The Ridge Regression was this time applied only to the training data. This allowed for the estimation of absolute and relative errors by applying the generated coefficients to the withheld test data.

The obtained model equations for the total data sets of lowC and highC isomers were the following:

lowC:

$$\mathrm{RI} = 858.25\mathrm{TDB4u} + 83.69nC - 7.99[\mathrm{CX4H3}] + 10.42[\mathrm{CX4H0}] - 2899.91 \quad (3)$$

highC:

$$\mathrm{RI} = -55.57\mathrm{TDB6u} + 123.41\mathrm{TDB7u} + 96.19nC - 40.53[\mathrm{CX4H3}] + 16.18[\mathrm{CX4H3}] - 407.36 \quad (4)$$

3. RESULTS AND DISCUSSION

3.1. Retention Index Model Performance. Due to the present uncertainties, calculated errors were always rounded up to the respective digit. The relative errors of both, the lowC and highC, models were less than 2% (1.3 and 1.6%, respectively). This resulted in mean absolute errors (MAE) of 13 RI (lowC) and 22 RI (highC), respectively, while presenting an overall MAE = 17 RI. The correlation of database and predicted RI for the combined highC and lowC models was apparent (R^2 = 0.9904, Figure 7). However, some

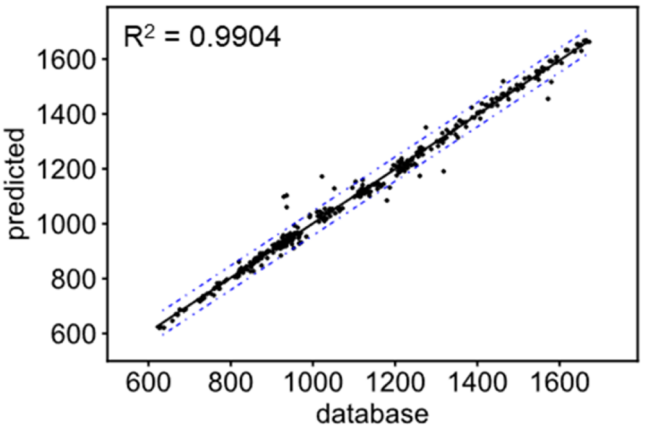


Figure 7. Overall RI model precision. Outliers were defined by ΔRI > 44, illustrated outside the blue dotted lines.

structural features were less precisely predictable and were considered to be outliers if the value of deviation between measured and predicted RI was double the MAE of the highC model (|ΔRI| > 44). This proved to be true for 14 out of 354 evaluated RI values, i.e., around 4% of the available isomers. Those outliers included especially cases with mixed types of branches, 3-methyl-5-propylnonane (ΔRI = −77); 3,4-diethyl-3,4-dimethylhexane (ΔRI = 55) or 3-ethyl-2,7-dimethyloctane (ΔRI = 95). With increasing complexity of structures, some patterns may have been underrepresented in literature data, possibly affecting the quality of prediction. Within 354 datapoints for *iso*-alkanes, approximately 60% consisted of molecules varying only in degree of methylation. The data set underrepresented simple (ethyl-, propyl-, or butyl-group) and

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

complex (multiple ethyl-, i-propyl-, or tert-butyl-group) structural motifs that contain longer side chains. Nevertheless, there were also several examples of precise prediction despite complex branch motifs (e.g., 7-ethyl-3-methylnonane; 3-ethyl-3,6-dimethyloctane; both ΔRI = 3). Furthermore, it has to be discussed that the literature data are not necessarily without outliers itself. Some structures were only once reported in the literature or did not have an assigned uncertainty to their measurement. Independent of the apparent uncertainties regarding the accuracy of the entire complexity of possible isomers, the model was considered sufficient for further procedure.

3.1.1. RI Prediction of iso-Alkane Data Set. The RI's of all 42,900 isomers[37] in the jet range[18,31,33] were calculated from their molecular descriptors, to complete the available data set. It was determined that 12 out of 42,900 isomer RI predictions were illogic outliers. Those cases presented the respective maximal branched isomers of *n*C13-*n*C17 (Figure 8), where no

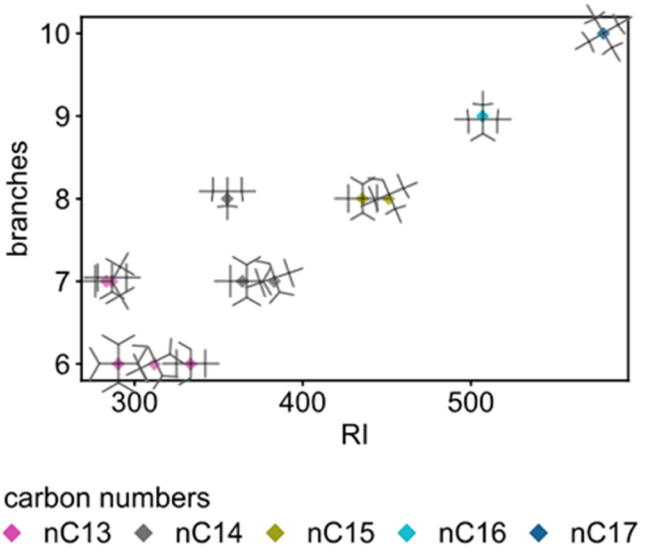


Figure 8. Physically illogic retention behavior (outliers) from the highC model predictions.

pathways of k = 7 were found between any atom pairs. Therefore, TDB7u was 0 in the respective molecule. For those cases, high deviations were introduced from typical elution behavior since the most important coefficient of the equation was neglected. These outliers were not considered for subgroup definition but were sorted in the highest respective available subgroup because of their compact nature. However, the occurrence of those structures in an actual fuel is expected to be improbable in general and, to the best of our knowledge, has not been reported yet.

3.2. Subgrouping of Isomeric Structures. Despite the quantification of the individual retention behavior, no single-component identification would be feasible. The accumulation of sometimes hundreds of isomeric structures within short retention ranges and the current uncertainty of prediction only enables an assignment of the individual isomers to certain retention areas. For this concept, precise retention ranges were therefore defined as a subgroup of the respective G-T and *n*C. The credibility of an isomer RI value was rated with decreasing priority from actual in-house calibrations over literature values (combined data set from NIST17) to model predictions. The RI's of highest available priority were then considered in this order for the substructure analysis and collectively sorted into the subsequently defined subgroups.

The proposed logic of structural subgrouping was created by the number of branches connected to the longest chain within each molecule. This strategy allowed a stepwise exclusion of specific isomers present at specific regions within the chromatogram. As depicted in Figure 9, isomers with a higher number of branches (*nB*) usually exhibited lower RI's, but also a higher range of possible retention. No distinct separation by the number of branches could be achieved since the high-branched isomers overlap into regions of smaller branching. On average the lower the *nB*, the smaller the range of RI, and the closer the RI of the respective structures to the linear *n*-alkane. The *n*-alkane always exhibits the highest RI of the alkane elution for the respective *n*C and was defined as the only structure in subgroup zero (SG0 = *n*-alkanes). The retention limit on the lower side of the isomeric trail of the individual *n*C was derived from the respective isomeric retention range (see Table 3). To create the range of RI that defined a subgroup, the lower RI boundary was defined close to the first eluting isomer of each group with an equal *nB*. The assigned subgroup number (SG#) refers to the number of branches it is defined by. Therefore, the SG# of each structure was typically, at maximum, equal to the *nB*. However, it was possible that isomers with higher *nB* were assigned a lower SG# if they exhibited a higher RI. The upper RI subgroup boundary on the other side was defined by the first isomer of

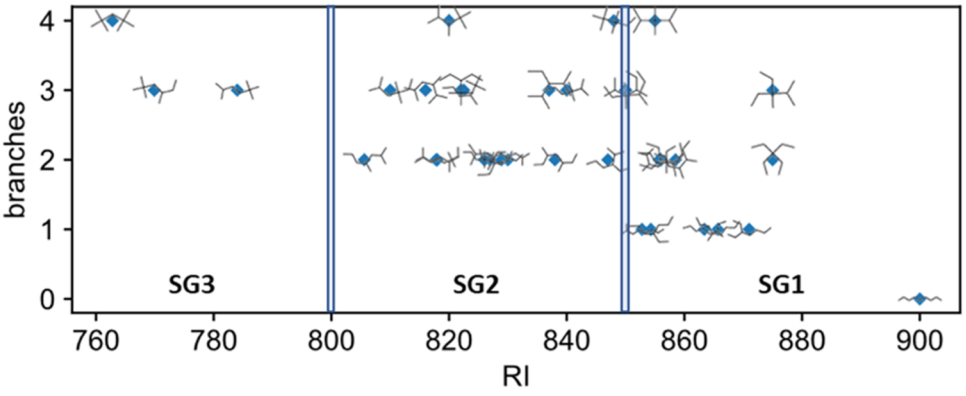


Figure 9. Subgroup definition illustrated on the retention behavior (by RI) of the 34 branched *iso*-nonanes by their number of branches per molecule (*nB*). Bold numbers represent the three respective subgroups defined for *n*C9 *iso*-alkanes. The *n*-alkane belongs to SG0.

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

Table 3. Defined *iso*-Alkane Subgroubs for the Jet Relevant Carbon Range *n*C7-*n*C17[a,b]

		subgroup data: lower RI boundary; database isomers/possible isomers; **mean branching**								
*n*C	possible isomers	1/1a	1/1b	2	3	4	5	6	7	$\overline{nB}_{nC}$
7	8	660		rest						
		4/4		4/4						
		1.25		**2.25**						**1.56**
8	17	740		700	rest					
		10/10		6/6	1/1					
		1.80		**2.50**	**3.00**					**2.00**
9	34	850		800	rest					
		12/12		19/19	3/3					
		1.83		**2.63**	**3.33**					**2.34**
10	74	940		900	865	rest				
		24/24		38/38	10/10	2/2				
		2.33		**2.84**	**3.30**	**3.50**				**2.72**
11	158	1040		1000	970	rest				
		10/18		14/91	1/39	3/10				
		1.72		**3.03**	**3.74**	**4.00**				**3.10**
12	354	1140		1100	1070	1030	rest			
		12/16		25/154	1/145	1/36	3/3			
		1.56		**3.03**	**3.89**	**4.58**	**4.33**			**3.48**
13	801	1250	1190	1100	1045	960	915	rest		
		8/8	29/43	2/278	1/263	1/173	2/20	0/11		
		1.38	**1.95**	**3.18**	**4.14**	**4.79**	**5.25**	**6.18**		**3.85**
14	1857	1340	1300	1200	1140	1070	1000	rest		
		10/10	11/40	7/487	0/652	0/533	0/116	0/19		
		1.20	**1.90**	**3.20**	**4.19**	**5.07**	**5.69**	**6.63**		**4.23**
15	4346	1440	1385	1290	1240	1180	1115	1070	rest	
		8/8	12/72	1/938	1/1208	0/1359	0/642	0/97	0/22	
		1.13	**1.94**	**3.36**	**4.32**	**5.13**	**5.84**	**6.62**	**7.36**	**4.61**
16	10,358	1540	1475	1385	1330	1270	1220	1160	rest	
		11/12	13/137	7/1697	0/2849	1/3404	0/1610	0/581	0/68	
		1.10	**2.18**	**3.52**	**4.50**	**5.36**	**6.11**	**6.81**	**7.78**	**4.99**
17	24,893	1640	1570	1490	1425	1375	1320	1260	rest	
		9/10	14/184	3/2241	0/6539	0/7185	0/5971	0/2407	0/356	
		1.10	**2.21**	**3.46**	**4.54**	**5.43**	**6.21**	**7.01**	**8.00**	**5.37**
sum	42,900	118/132	79/476	126/5953	18/11714	8/12702	5/8362	0/3115	0/466	

[a]Information is listed by carbon number (*n*C) and respective subgroup. [b]First line: the lower RI boundary of the subgroup. Second line: the number of classified database isomers compared to the number of theoretic isomers in the subgroup (Database/Total). Third line: the **average number of branches** ($\overline{nB}$) over all *iso*-alkanes within the subgroup. Additionally, the mean number of branches for every *n*C ($\overline{nB}_{nC}$) is given. Secondary 1a and 1b subgrouping only available for *n*C13-*n*C17.

the next lowest *nB* or for SG1 defined by the *n*-alkane of the same *n*C. With respect to the prediction uncertainties, subgroup sizes for the lowC-model region were set to a minimum width of 30 RI, and for the highC-model region, not defined smaller than 40 RI.

For example, for *iso*-nonane (Figure 9), the lower boundary of SG2 was set to RI = 800, since 2,4-dimethylheptane was the first eluting isomer with *nB* = 2 at RI = 806. In contrast, 2,2-diethylpentane has two branches (*nB* = 2) but was found in subgroup 1 (RI > 850) because its RI (875) was higher than the RI of the first eluting isomer with *nB* = 1 (2-methyloctane; RI = 853). With increasing *n*C the possible number of subgroups increased naturally. However, small differences between the elution ends of the highest branching groups were considered insignificant and were combined in the respective "last" bin of each *n*C (e.g., no SG4 in Figure 9 and no SG7, or higher, in Figure 10). Simplified, the SG# can be interpreted as

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

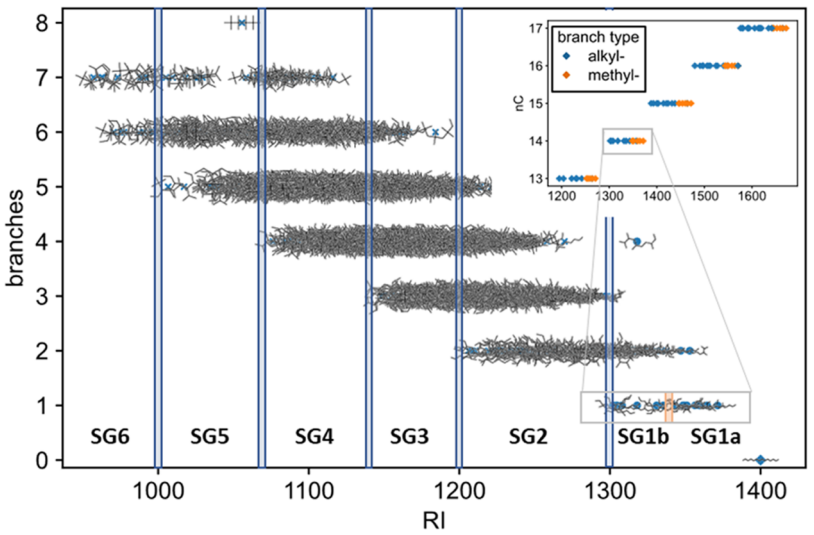


Figure 10. Depiction of isomeric subgroups (1−6) for the *n*C14 *iso*-alkanes. Added illustration on possible branch differentiation for single-branched isomers (1a, methylated; 1b, alkylated) in the highC model range (upper right). The *n*-alkane belongs to SG0.

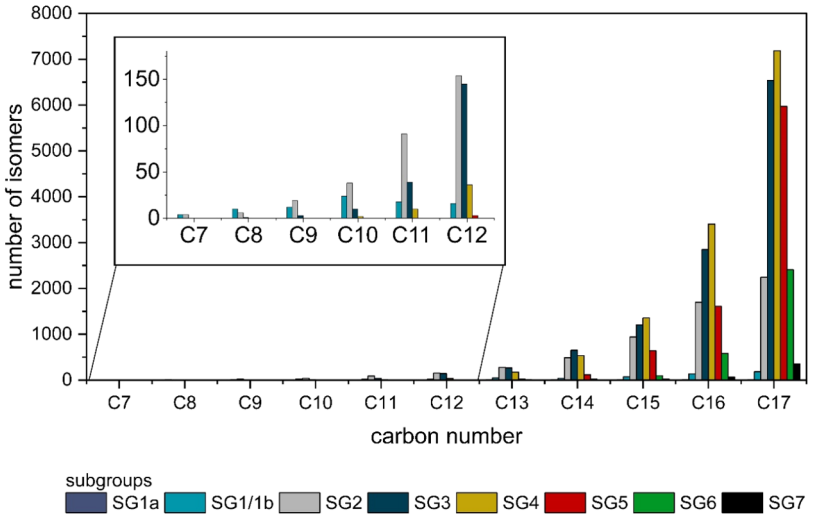


Figure 11. Illustration of carbon number-dependent rise of structure possibilities and change in distribution within the respective subgroups. Numbers of the *n*C7-*n*C12 section are magnified for visualization. Subgroup 1 is separated into 1a and 1b for the highC isomers (*n*C13-*n*C17).

only isomers with $nB \geq$ SG# occurring in the respective subgroup.

With higher *n*C the variety and impact of the branch type also became partially considerable. An evaluation on alkyl vs methyl chains in the monobranched isomers implied that the type of carbon chain could also be resolved. In general, the longer the carbon chain of the individual branch, the lower the RI. Due to this behavior, it was possible to introduce a further subclassification. This was possible for the first subgroup (SG1) for the highC modeled isomers. Further differentiation between all monomethylated (1a) isomers from the mono-alkylated (e.g., ethylated-, propylated-, etc.) structures (1b) was implemented above *n*C12. This usually resulted in an almost complete separation, only 3- and 4-ethyl-branched isomers were often found to be overlapping into the SG1a, while isomers with propyl- or higher chains eluted noticeably earlier. Figure 10 provides an additional illustration of the subgrouping strategy for higher structural diversity with all *n*C14 isomers. It furthermore highlights the mentioned separation of branch length within the highC model region in SG1. All of the subgroup definitions as well as the respective number of assigned isomers can be found in Table 3.

As is visible for the *n*C14 isomers in Figure 10, some ranges of the retention space were crowded by relatively high numbers of isomers. Thus, the isomers were not evenly distributed throughout the subgroups. For the higher *n*C, this resulted in a Gaussian-like distribution around the respective central subgroup (Figure 11). The available data from measurements or the literature data were heavily focused on the lower subgroup numbers. The higher the subgroup number, the lower the (relative) number of datapoints. From SG4 on, only 13 reference RI were obtainable for all included *n*C. The SG6 and SG7 had even no reference data at all. On the contrary, especially single-branched isomers were well represented in the data set (110/119). Thus, uncertainties for retention behavior of especially heavier branched isomers remain to be further investigated.

3.3. Definition of Branching Indicators. For convenient interpretation of the *iso*-alkane subgroups and their application to fuel samples, a number of branching indicators were

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

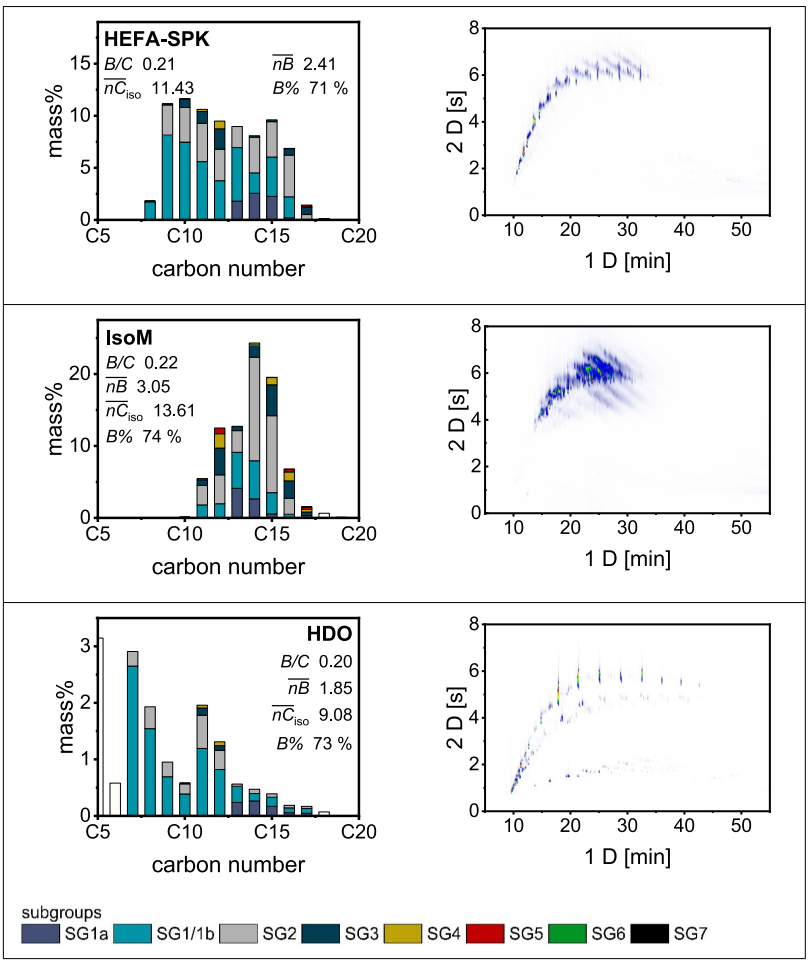


Figure 12. *iso*-Alkane subgroup distributions (carbon number vs mass%) in HEFA-SPK, IsoM, and HDO fuel examples, and their respective GCxGC chromatograms. Retention time on first dimension (1D) against retention time on the second dimension (2D).

additionally defined. For each subgroup (SG) of each carbon number (nC) the mean number of branches ($\overline{nB}_{SG(nC)}$, Table 3) was calculated by normalization of the number of branches (nB) against the number of isomers (eq 5).

$$\overline{nB}_{SG(nC)} = \frac{\sum_{i=1}^{n} nB_i}{n} \tag{5}$$

where i is an individual isomer within the respective subgroup of a distinct nC. With the quantified molar percentages (mol%) of each SG from GCxGC, the mean number of branches ($\overline{nB}$) for the total isoparaffinic fraction ($\text{mol\%}_{iso}$) of the sample was derived.

$$\overline{nB} = \frac{\sum_{nC=7}^{17} \sum_{i=1}^{n} \overline{nB}_{SG_i(nC)} \times \text{mol\%}_{SG_i(nC)}}{\text{mol\%}_{iso}} \tag{6}$$

To contextualize the $\overline{nB}$ to different carbon distributions of the respective samples, the mean number of branches per carbon (B/C) was furthermore calculated. This was done by normalization with the mean carbon number within the considered isoparaffinic fraction ($\overline{nC}_{iso}$).

$$B/C = \frac{\overline{nB}}{\overline{nC}_{iso}} \tag{7}$$

To analyze the influence of subgrouping in contrast to total uniform isomeric distribution, the mean number of branches for all isomers of each carbon number was determined as well ($\overline{nB}_{nC}$). The calculation was equal to that in eq 6, with i being the individual isomer within each nC, without subgroup differentiation. This way, it was possible to assess the effect of subgroup detail over uniform distribution. It describes whether a sample exhibits above or below uniform branching, contextualized in percent ($B\%$).

$$B\% = \frac{\overline{nB}}{\overline{nB}_{nC}} \tag{8}$$

Since the model approach focused only on *iso*-alkanes within $nC7$-$nC17$, the average carbon number of isomers within the model region ($\overline{nC}_{iso}$) was calculated separately to the overall mean carbon number of the sample ($\overline{nC}_{total}$). The result of the average branching could then be evaluated in comparison with the general hydrocarbon cut of a sample or in connection with existing *iso*-alkane fractions outside the carbon range of the models. If both mean carbon numbers were sufficiently close to one another, effective description of the total *iso*-alkane branching was assumed.

3.4. Application of Subgroups and Branching Indicators on Fuel Samples. The effect of subgroup quantification on property predictions for fuel assessment is beyond the scope of the present work and will be addressed separately. Nevertheless, potential effects could already be evaluated by considering the distribution of certain properties

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

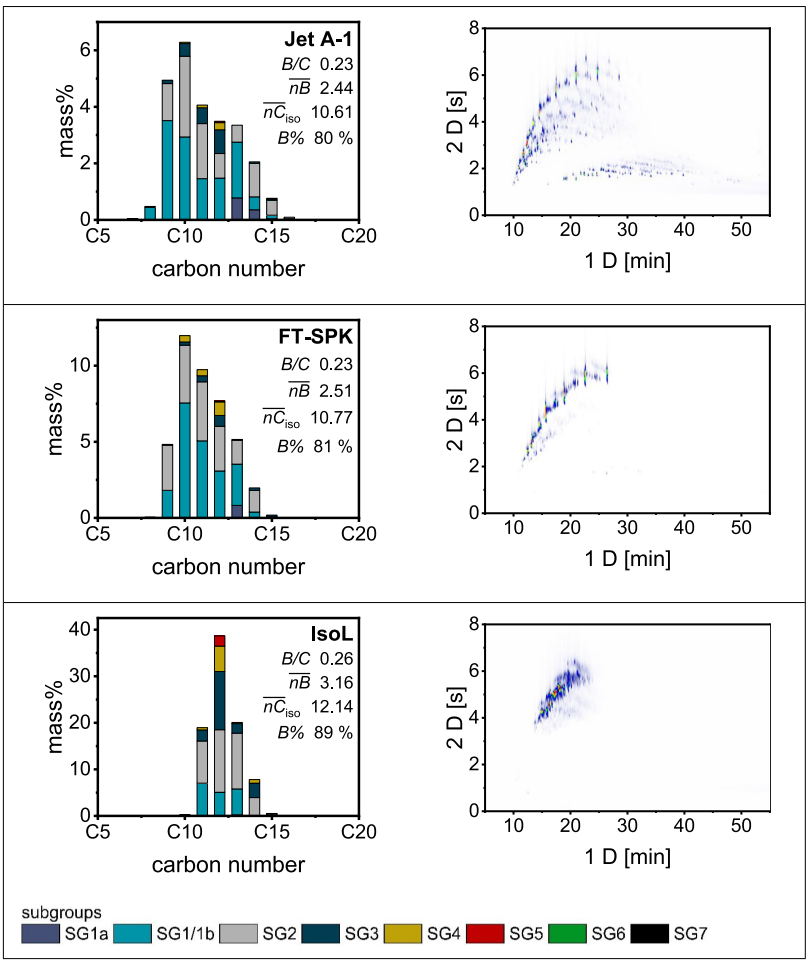


Figure 13. *iso*-Alkane subgroup distributions (carbon number vs mass%) in Jet A-1, FT-SPK, and IsoL fuel examples, and their respective GCxGC chromatograms. Retention time on first dimension (1D) against retention time on the second dimension (2D).

throughout different subgroups (see the SI) and the actual isomeric differences in investigated example fuels. To test the herein created *iso*-alkane subgroup strategy, eight fuel samples of different origins were measured in the GCxGC-FID/qMS setup. The investigated fuels contained three fuels from fossil origin and five fuels from synthetic processes. One conventional reference kerosene (Jet A-1) and two commercially available isoparaffinic solvents (ISOPAR L, ISOPAR M) from fossil origin were analyzed. One HEFA-SPK, one ATJ-SPK, one FT-SPK, one hydro-deoxygenated Fuel (HDO), and one intermediate crude product from methanol-to-jet model reactions (MTJ crude) were selected to cover synthetic fuel pathways. Coelution of *iso*-alkanes with complex *iso*-alkenes can induce problems within smaller subgroups but is neglectable for prescreening of hydrotreated jet fuel candidates. The presence of mass fragments −2 H can be utilized to determine the extent of overlap if coelution is in question for unfinished products. This was solely necessary for the assessment of the MTJ Crude sample.

Although GCxGC presents a significant improvement in overall resolution and peak capacity over its one-dimensional counterpart, the resolution in the first dimension is dependent on the modulation procedure. In the scope of this article, the modulation time of 8 s led to the 1D pixel resolution of several RI. Despite achieving a good modulation of low-mass isomers, insufficient chromatographic 1D separation resulted in "crowded" signals among low *n*Cs. Exemplary, between *n*-heptane and *n*-octane, only 6 modulations were possible, which resulted in average RI steps of 16.67 per modulation. However, the low resolution was compensated by single-compound calibration. For heavier *iso*-alkanes, the uncertainty of the retention behavior was the greater challenge for correct subgroup assignment, since the chromatographic resolution improved significantly to a constant 4 ± 0.5 RI per modulation for $nC > 11$. Nevertheless, the actual defined boundaries of specific subgroup RI's could not always fit perfectly to the defined value. In cases where the boundary was laid within a modulation period, the subgroup boundaries were created at the end of the prior modulation. This way, by definition, the risk of overlapping isomer signals from a smaller bin was consistently reduced since they might already have been present within the modulation period. Exceptions to this rule applied if only the second decimal digit of the detected retention time was deviating from the calculated retention time of the defined RI boundary. (See the Supporting Information for exemplary application of subgroups within the example fuels. In order to make the retention areas of the subgroups more visible, tailing and signal intensities below the S/N ratio were graphically minimized. Therefore, the areas were only roughly drawn in the graphics to primarily portray the typical areas within the chromatogram.)

In general, the eight fuels exhibited different carbon ranges and forms of distribution. Solely focusing on the G-T of *iso*-alkanes (Figures 12−14), the FT-SPK, the isoparaffinic

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

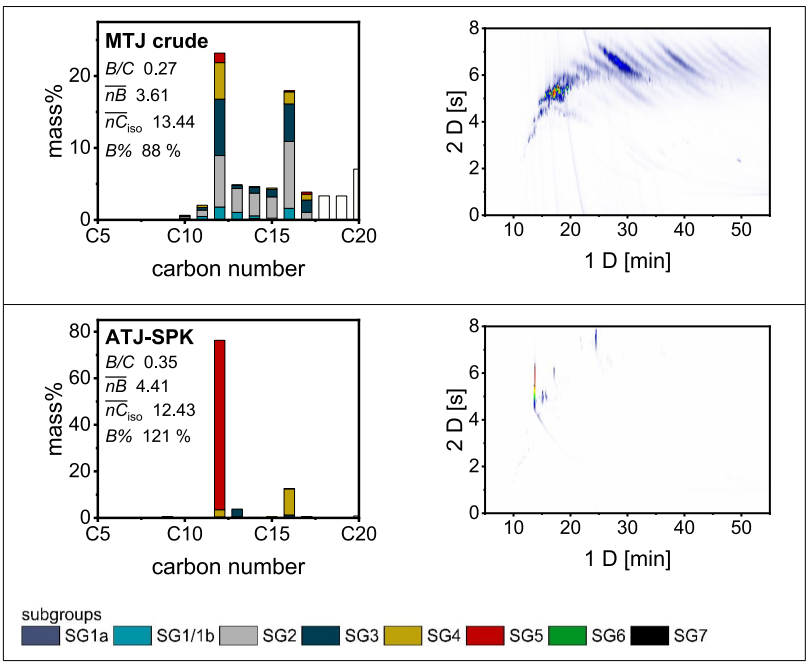


Figure 14. *iso*-Alkane subgroup distributions (carbon number vs mass%) in an MTJ crude and an ATJ-SPK fuel example and their respective GCxGC chromatograms. Retention time on first dimension (1D) against retention time on the second dimension (2D).

solvents (IsoL and IsoM), and the Jet A-1 fuel showed a Gaussian distribution around one central carbon. The *iso*-alkanes in the HEFA-SPK fuel were more evenly distributed throughout the different *n*C (see *iso*-alkane distribution throughout the different *n*C in Figures 12 and 13). The ATJ-SPK and MTJ Crude were found to be visibly different from the rest as they were synthetically built with a focus on C4-building blocks, which led to high concentrations within the *n*C12, *n*C16, etc. isomers (Figure 14). Despite the similar distribution to the carbon numbers, the isomeric diversity was evidently much higher in the MTJ Crude over the ATJ-SPK, which is already visible from the chromatograms as well. In the GCxGC graphics (Figure 14) the ATJ-SPK exhibited almost no further structures than two specific isomer signals of high intensity, which were already reported in the literature.[18,54] The MTJ crude process on the contrary showed a large distribution of signals around a centrally positioned isomer in the chromatogram. The diverse distribution of individual structures was also visible from the relatively long *iso*-alkane signal tracks, compared to the other fuel examples. Lastly, HDO provided an individual carbon range distribution with two local maxima, with especially high *iso*-alkane numbers in the low-carbon (*n*C7) and medium-carbon (*n*C11) range.

The amount of branching in the test samples was indicated to be limited, as the first and second subgroups usually dominated the quantification. This corresponded primarily to single-, double-, and sometimes triple-branched *iso*-alkanes. Especially in HEFA-SPK, HDO, Jet A-1, and FT-SPK, those two subgroups accounted for mass percentages above 89.9% (see ratios Figures 12−14 and Table 4). In comparison, a greater range of subgroups and subsequently higher mass percentages of subgroup 3, and higher, were more prominently found in IsoM (mass% SG3+ = 24.7%), IsoL (mass% SG3+ = 34.4%), and the MTJ crude (mass% SG3+ = 44.7%). This provided already a hint that the isomers present in those fuels probably had a higher mean number of branches. The ATJ-SPK was the only fuel example that essentially consisted only of highly branched isomers (mass% SG3+ = 97.7%). Nevertheless, it should be noted that subgroup numbers alone cannot define specific branching due to the possibility of structural overlapping. The sample size of evaluated fuels was too small to identify clear indications for *iso*-alkane distribution differences based on fuel feedstock or production pathway.

Table 4. Summary of Branching Indicators of the Analyzed Fuel Samples[a]

	$\overline{nB}$	$\overline{nC}_{iso}$	*B*/*C*	$\overline{nB}_{nC}$	m% (SG1 + SG2)	*B*%
HEFA-SPK	2.41	11.43	0.21	3.37	91.2%	71%
IsoM	3.05	13.61	0.22	4.10	74.5%	74%
HDO	1.85	9.08	0.20	2.44	97.0%	73%
Jet A-1 REF	2.44	10.61	0.23	3.03	89.9%	80%
FT-SPK	2.51	10.77	0.23	3.06	91.6%	81%
IsoL	3.16	12.14	0.26	3.53	65.6%	89%
MTJ crude	3.61	13.44	0.27	4.08	55.3%	88%
ATJ-SPK	4.41	12.43	0.36	3.65	1.4%	121%

[a]The table lists the following indicators within the model range *n*C7-*n*C17: the mean branching number ($\overline{nB}$) through subgroup consideration, the mean carbon number of *iso*-alkanes ($\overline{nC}_{iso}$) within the model region, the branches per carbon atom (*B*/*C*) within the model region, the fuels mean number of branches with uniform *iso*-alkane distribution ($\overline{nB}_{nC}$), the summarized mass percentages of subgroup 1 and subgroup 2 m% (SG1+SG2) from GCxGC results, and the fuels mean subgroup branching relative to mean uniform isomer distribution (*B*%).

Depending on the respective examined branching indicator, a slightly different interpretation of the fuels was possible. The mean number of branches ($\overline{nB}$) for example could paint a different picture, if no normalization against the carbon distribution was conducted. It could introduce a bias to the evaluation of high-carbon fuels that have naturally more branching possibilities but fewer branching points relative to the total molecule size. This could be observed for IsoM,

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

where the average molecule was theoretically exhibiting 3.05 branches, which was among the highest $\overline{nB}$ found. Yet, similar relative branching (B%) would allow more branching points since the sample exhibited a fairly high carbon distribution ($\overline{nC}_{iso}$) in contrast to the other low B% fuel examples (Table 4).

In contrast to $\overline{nB}$, the branches per carbon (B/C) and the B% were found to be well suited as general indicators of branching without further sample context. Through the evaluation of both indicators, it was apparent that the IsoM fuels contained almost the lowest relative branching (B/C = 0.22; B% = 74%). Nevertheless, all three indicators could provide relevant context for understanding the respective fuels. Since fuel assessment is dependent on specific structures, the $\overline{nB}$ could still be helpful independently of the relative branching information.

The overall smallest $\overline{nB}$ = 1.85 was found in the HDO fuel. Since the total *iso*-alkane content was significantly influenced by the amount of low-carbon isomers, the branching possibility was also limited in general. The average branching per carbon and the relative branching percentage (B%) were comparable to those of HEFA-SPK and IsoM samples.

When comparing the conventional Jet Fuel with the synthetic FT-SPK product, the isomeric distribution was found to be almost similar in all branching indicators (Table 4). In Jet A-1, FT-SPK, as well as HEFA-SPK and HDO, three isomer signals seemed to mainly influence the overall *iso*-alkane composition of each carbon number. These signals were identified as likely being the 2/3/4-monomethylated alkanes (according to NIST MS reference data and according to the available RI values). Therefore, the average branching would be highly influenced by the high occurrence of single-branched isomers. The approximation of equal distribution within the subgroup thus has to be considered when evaluating the precision of the calculated indicators.

The investigated examples for ATJ-SPK and MTJ proved to show higher branching by subgroup detail, chromatogram, and through the defined indicators. Especially the MTJ crude showed no apparent preferences of isomeric structure but high diversity, which was apparent from the aforementioned long isomeric tracks in the chromatogram. Moreover, it looked almost evenly distributed around the maximum between subgroups 2 to 4. The isoparaffinic solvent IsoL seemed to have similar results of relative branching. Although the nC maximum was higher, the distillation cut was narrower and consequently isomers from higher carbons were excluded. This led to a smaller average carbon number and subsequently less possible branching. The ATJ-SPK fuel composition, as reported, is dominated by 2,2,4,6,6-pentamethylheptane (>70 mass%) and 2,2,4,4,6,8,8-heptamethylnonane (>10 mass %).[18,54] Both signals were correctly identified as within the expected subgroups. However, the fact that other signals of lower nB occur within their subgroup, the calculated average branching even underestimated the actual amount of branching. The uncertainty of the indicators has to be considered again for such fuels with prominent individual signals since the new subgroup detail does not resolve the subgroups beyond equal distribution. Yet, they clearly illustrated the expected result as the ATJ-SPK was the only fuel investigated with branching above the average equal distribution (B% = 121%).

The first application of subgroup boundaries on the GCxGC results seemed overall promising. The small number of samples already proved to allow differentiation of the isomeric occurrence and the deficiencies of simplifying the *iso*-alkane compounds into groups only by nC.

4. CONCLUSIONS AND OUTLOOK

A typical composition-based fuel assessment usually only considers hydrocarbon group-type (G-T) and carbon number (nC)-dependent GCxGC results. The structural information on specific isomers is generally not provided, as the reference data is often not sufficient to evaluate the highly structurally diverse hydrocarbon groups individually. This limitation induces higher uncertainties for the prediction of some essential fuel properties (e.g., flash point, freezing point, viscosity). To generate information about the structural distribution of *iso*-alkanes in novel fuels, a structure-to-retention investigation was carried out. This was done by a multilinear regression model approach to predict missing retention indices (RI) for all 42,900 possible isomer structures, within *n*C7-*n*C17 *iso*-alkanes. The completed RI database from measurements, literature data, and predictions was subsequently used to create structural-dependent subgroups throughout the retention space for each nC. Subgroup classifications were utilized to derive quantitative branching indicators for comparison of isomeric structure and distributions within different fuels. It was confirmed that the analyzed fuel samples contained varying isomeric distributions throughout the subgroups, which proves the necessity of substructural information in fuel characterization and assessment. The investigated ATJ-SPK and MTJ crude fuel samples were found to be dominated by isomers with high branching. Higher mean branching was both obtainable through singular structures and high diversity throughout the subgroups. On the contrary, the fuel samples of HEFA-SPK, FT-SPK, and HDO were shown to exhibit similar isomeric distribution to the conventional Jet A-1 fuel sample, mainly categorized as low- or single-branched isomers. The subgroup information is expected to improve property and reactivity predictions for fuels by enhancing structural identification and understanding the presence or absence of specific isomer groups. The effects of this new subgroup detail on total prescreening results will need further systematic investigation of fuel samples. In order to expand the structural detail to more diverse sample compositions, which are not dominated by the *iso*-alkane fraction, the development of subgroups for other hydrocarbon groups (e.g., *cyclo*-alkanes) will be a future target.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.energyfuels.4c05573.

Discussion of subgroup impact on different properties, including structural RI-to-property illustrations; representation of exemplary subgroup assignment according to retention behavior of the eight discussed example fuels (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Hannes Lüdtke − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany;*

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

orcid.org/0009-0002-4772-8245;
Email: Hannes.Luedtke@dlr.de

Authors

Florian Pütz − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany;* orcid.org/0009-0006-5270-2030

Samuel Grams − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany*

Uwe Bauder − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany;* orcid.org/0000-0002-5019-6043

Markus Köhler − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany;* orcid.org/0000-0001-9562-8455

Andreas Huber − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany*

Patrick Oßwald − *DLR, German Aerospace Center, Institute of Combustion Technology, 70569 Stuttgart, Germany;* orcid.org/0000-0002-2257-2988

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.energyfuels.4c05573

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge funding by the German Federal Ministry for Digital and Transport (BMDV) within the SAFari project (FKZ: 16RK14009E). Furthermore, the support by DLR's Scientific Program Administrations Energy, Aeronautics, Space and Transport in the project framework of DLR NeoFuels is gratefully acknowledged.

■ ABBREVIATIONS

ASTM	American Society for Testing and Materials
ATJ	alcohol to jet
B%	subgroup branching/uniform branching [%]
B/*C*	branches per carbon (in the average isomer)
FID	flame ionization detector
FT	Fischer−Tropsch
GCxGC	two-dimensional gas chromatography
G-T	group-type
HEFA	hydroprocessed esters and fatty acids
highC	high-carbon model
ID	inner diameter
lowC	low-carbon model
nB	number of branches
$\overline{nB}_{SG}$	mean number of branches per molecule within a subgroup
$\overline{nB}_{nC}$	mean number of branches per molecule within the same carbon number
$\overline{nB}$	mean number of branches per molecule within a sample (based on subgroup detail)
$\overline{nC}_{iso}$	mean carbon number (modeled *iso*-alkanes only)
$\overline{nC}$	mean carbon number (of the total sample)
nC	carbon number
qMS	quadrupole mass spectrometer
RI	retention index
SAF	sustainable aviation fuel
SG	subgroup
SPK	synthetic paraffinic kerosene

■ REFERENCES

(1) Lee, D. S.; Fahey, D. W.; Skowron, A.; Allen, M. R.; Burkhardt, U.; Chen, Q.; Doherty, S. J.; Freeman, S.; Forster, P. M.; Fuglestvedt, J.; Gettelman, A.; De León, R. R.; Lim, L. L.; Lund, M. T.; Millar, R. J.; Owen, B.; Penner, J. E.; Pitari, G.; Prather, M. J.; Sausen, R.; Wilcox, L. J. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmos. Environ.* **2021**, *244*, No. 117834.

(2) Teoh, R.; Schumann, U.; Gryspeerdt, E.; Shapiro, M.; Molloy, J.; Koudis, G.; Voigt, C.; Stettler, M. E. J. Aviation contrail climate effects in the North Atlantic from 2016 to 2021. *Atmos. Chem. Phys.* **2022**, *22* (16), 10919−10935.

(3) Teoh, R.; Schumann, U.; Voigt, C.; Schripp, T.; Shapiro, M.; Engberg, Z.; Molloy, J.; Koudis, G.; Stettler, M. E. J. Targeted Use of Sustainable Aviation Fuel to Maximize Climate Benefits. *Environ. Sci. Technol.* **2022**, *56* (23), 17246−17255.

(4) Voigt, C.; Kleine, J.; Sauer, D.; Moore, R. H.; Bräuer, T.; Le Clercq, P.; Kaufmann, S.; Scheibe, M.; Jurkat-Witschas, T.; Aigner, M.; Bauder, U.; Boose, Y.; Borrmann, S.; Crosbie, E.; Diskin, G. S.; DiGangi, J.; Hahn, V.; Heckl, C.; Huber, F.; Nowak, J. B.; Rapp, M.; Rauch, B.; Robinson, C.; Schripp, T.; Shook, M.; Winstead, E.; Ziemba, L.; Schlager, H.; Anderson, B. E. Cleaner burning aviation fuels can reduce contrail cloudiness. *Commun. Earth Environ.* **2021**, *2* (1), 114.

(5) Gierens, K.; Sausen, R.; Bauder, U.; Eckel, G.; Großmann, K.; Le Clercq, P.; Lee, D. S.; Rauch, B. *Influence of Aviation Fuel Composition on the Formation and Lifetime of Contrails — A Literature Review*, Research Report; Concawe, 2024.

(6) Schripp, T.; Grein, T.; Zinsmeister, J.; Oßwald, P.; Köhler, M.; Müller-Langer, F.; Hauschild, S.; Marquardt, C.; Scheuermann, S.; Zschocke, A.; Posselt, D. Technical application of a ternary alternative jet fuel blend − Chemical characterization and impact on jet engine particle emission. *Fuel* **2021**, *288*, No. 119606.

(7) European Parliament. *Verordnung (EU) 2023/2405 des Europäischen Parlaments und des Rates vom 18. Oktober 2023 zur Gewährleistung gleicher Wettbewerbsbedingungen für einen nachhaltigen Luftverkehr (Initiative ReFuelEU Aviation)*. http://data.europa.eu/eli/reg/2023/2405/oj (accessed Sept 25, 2024).

(8) ICAO through the ICAO-UNDP Project, financed by GEF. *Sustainable Aviation Fuels Guide: Version 2*. https://www.icao.int/environmental-protection/Documents/Sustainable%20Aviation%20Fuels%20Guide_100519.pdf (accessed Oct 25, 2024).

(9) ICAO. *Initiatives and Projects*. https://www.icao.int/environmental-protection/GFAAF/Pages/InitiativesAndProjects.aspx (accessed Oct 25, 2024).

(10) NOW GmbH. *SAF-Monitor: Globally Announced Production Capacities for Sustainable Aviation Fuels*. https://erneuerbarekraftstoffe.de/saf-monitor/ (accessed Oct 25, 2024).

(11) Undavalli, V.; Olatunde, O. B. G.; Boylu, R.; Wei, C.; Haeker, J.; Hamilton, J.; Khandelwal, B. Recent advancements in sustainable aviation fuels. *Prog. Aerosp. Sci.* **2023**, *136*, No. 100876.

(12) Märkl, R. S.; Voigt, C.; Sauer, D.; Dischl, R. K.; Kaufmann, S.; Harlaß, T.; Hahn, V.; Roiger, A.; Weiß-Rehm, C.; Burkhardt, U.; Schumann, U.; Marsing, A.; Scheibe, M.; Dörnbrack, A.; Renard, C.; Gauthier, M.; Swann, P.; Madden, P.; Luff, D.; Sallinen, R.; Schripp, T.; Le Clercq, P. Powering aircraft with 100% sustainable aviation fuel reduces ice crystals in contrails. *Atmos. Chem. Phys.* **2024**, *24* (6), 3813−3837.

(13) Quante, G.; Voigt, C.; Kaltschmitt, M. Targeted use of paraffinic kerosene: Potentials and implications. *Atmos. Environ.: X* **2024**, *23*, No. 100279.

(14) Schripp, T.; Herrmann, F.; Oßwald, P.; Köhler, M.; Zschocke, A.; Weigelt, D.; Mroch, M.; Werner-Spatz, C. Particle emissions of two unblended alternative jet fuels in a full scale jet engine. *Fuel* **2019**, *256*, No. 115903.

(15) Hall, C.; Bell, D. C.; Feldhausen, J.; Rauch, B.; Heyne, J. Quantifying isomeric effects: A key factor in aviation fuel assessment and design. *Fuel* **2024**, *357*, No. 129912.

(16) Kathrotia, T.; Oßwald, P.; Zinsmeister, J.; Methling, T.; Köhler, M. Combustion kinetics of alternative jet fuels, Part-III: Fuel modeling and surrogate strategy. *Fuel* **2021**, *302*, No. 120737.
(17) Kathrotia, T.; Oßwald, P.; Naumann, C.; Richter, S.; Köhler, M. Combustion kinetics of alternative jet fuels, Part-II: Reaction model for fuel surrogate. *Fuel* **2021**, *302*, No. 120736.
(18) Oßwald, P.; Zinsmeister, J.; Kathrotia, T.; Alves-Fortunato, M.; Burger, V.; van der Westhuizen, R.; Viljoen, C.; Lehto, K.; Sallinen, R.; Sandberg, K.; Aigner, M.; Le Clercq, P.; Köhler, M. Combustion kinetics of alternative jet fuels, Part-I: Experimental flow reactor study. *Fuel* **2021**, *302*, No. 120735.
(19) Oh, J.; Oldani, A.; Lee, T.; Shafer, L. In *Deep Learning Algorithms for Assessing Sustainable Jet Fuels from Two-Dimensional Gas Chromatography*, Proceedings of the AIAA SciTech Forum: AIAA Science and Technology Forum and Exposition, San Diego, USA, January 3−7; American Institute of Aeronautics and Astronautics Inc, 2022.
(20) Vozka, P.; Modereger, B. A.; Park, A. C.; Zhang, W. T. J.; Trice, R. W.; Kenttämaa, H. I.; Kilaz, G. Jet fuel density via GC × GC-FID. *Fuel* **2019**, *235*, 1052−1060.
(21) Hall, C.; Creton, B.; Rauch, B.; Bauder, U.; Aigner, M. Probabilistic Mean Quantitative Structure−Property Relationship Modeling of Jet Fuel Properties. *Energy Fuels* **2022**, *36* (1), 463−479.
(22) Hall, C.; Rauch, B.; Bauder, U.; Le Clercq, P.; Aigner, M. In *Machine Learning Algorithms for Fuel Property Prediction in Comparison with State of the Art Physical Models*, Proceedings of the 16th International Conference on Stability Handling and Use of Liquid Fuels: September 8−12, 2019; Morris, R. E., Ed.; Nova Research, Inc.: Long Beach, USA, 2019.
(23) Hall, C.; Rauch, B.; Bauder, U.; Aigner, M. In *Comparison of Probabilistic Jet Fuel Property Models for the Fuel Screening and Design*, Proceedings of the 17th International Conference on Stability, Handling and Use of Liquid Fuels: September 11−15, 2022; Morris, R. E., Ed.; Curran Associates Inc: Dresden, Germany, 2023.
(24) Shi, X.; Li, H.; Song, Z.; Zhang, X.; Liu, G. Quantitative composition-property relationship of aviation hydrocarbon fuel based on comprehensive two-dimensional gas chromatography with mass spectrometry and flame ionization detector. *Fuel* **2017**, *200*, 395−406.
(25) Feldhausen, J.; Bell, D. C.; Yang, Z.; Faulhaber, C.; Boehm, R.; Heyne, J. Synthetic aromatic kerosene property prediction improvements with isomer specific characterization via GCxGC and vacuum ultraviolet spectroscopy. *Fuel* **2022**, *326*, No. 125002.
(26) Kosir, S.; Landsaw, A.; Steinecker, W. H.; West, Z. J.; Zabarnick, S. Synthetic Blend Component Study: The Effects of Hydrocarbon Composition on Aviation Fuel Dielectric Constant. *Energy Fuels* **2024**, *38* (17), 16451−16457.
(27) Heyne, J.; Rauch, B.; Le Clercq, P.; Colket, M. Sustainable aviation fuel prescreening tools and procedures. *Fuel* **2021**, *290*, No. 120004.
(28) Bauder, U.; Hall, C.; Pütz, F.; Enderle, B.; Le Clercq, P. C.; Rauch, B. In *An Interactive, Interdisciplinary and Collaborative Digital Platform for the Assessment and Optimization of Jet Fuels*, Proceedings of the 17th International Conference on Stability, Handling and Use of Liquid Fuels: September 11−15, 2022; Morris, R. E., Ed.; Curran Associates Inc: Dresden, Germany, 2023.
(29) Yang, Z.; Kosir, S.; Stachler, R.; Shafer, L.; Anderson, C.; Heyne, J. S. A GC × GC Tier α combustor operability prescreening method for sustainable aviation fuel candidates. *Fuel* **2021**, *292*, No. 120345.
(30) Mostafa, A.; Edwards, M.; Górecki, T. Optimization aspects of comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* **2012**, *1255*, 38−55.
(31) Heyne, J.; Bell, D.; Feldhausen, J.; Yang, Z.; Boehm, R. Towards fuel composition and properties from Two-dimensional gas chromatography with flame ionization and vacuum ultraviolet spectroscopy. *Fuel* **2022**, *312*, No. 122709.
(32) Jennerwein, M. K.; Eschner, M.; Gröger, T.; Wilharm, T.; Zimmermann, R. Complete Group-Type Quantification of Petroleum Middle Distillates Based on Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry (GC × GC-TOFMS) and Visual Basic Scripting. *Energy Fuels* **2014**, *28* (9), 5670−5681.
(33) Melder, J.; Zinsmeister, J.; Grein, T.; Jürgens, S.; Köhler, M.; Oßwald, P. Comprehensive two-dimensional gas chromatography: a universal method for composition-based prediction of emission characteristics of complex fuels. *Energy Fuels* **2023**, *37* (6), 4580−4595.
(34) Striebich, R. C.; Shafer, L. M.; Adams, R. K.; West, Z. J.; DeWitt, M. J.; Zabarnick, S. Hydrocarbon Group-Type Analysis of Petroleum-Derived and Synthetic Fuels Using Two-Dimensional Gas Chromatography. *Energy Fuels* **2014**, *28* (9), 5696−5706.
(35) Ettre, L. S. Nomenclature for chromatography: (IUPAC Recommendations 1993). *Pure Appl. Chem.* **1993**, *65* (4), 819−872.
(36) National Institute of Standards and Technology. *NIST 17 MS Database and MS Search Programm v.2.3*; U.S. Department of Commerce, 2017.
(37) Gugisch, R.; Kerber, A.; Kohnert, A.; Laue, R.; Meringer, M.; Rücker, C.; Wassermann, A. MOLGEN 5.0, A Molecular Structure Generator. In *Advances in Mathematical Chemistry and Applications*; Restrepo, G.; Villaveces, J. L.; Basak, S. C., Eds.; Bentham Science Publishers, 2014; Vol. *1*, pp 113−138.
(38) Kováts, E. Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentionsindices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone. *Helv. Chim. Acta* **1958**, *41* (7), 1915−1932.
(39) Dimov, N. General model for precalculation of the retention indices of isoalkanes separated by gas or liquid chromatography. *Anal. Chim. Acta* **1987**, *201*, 217−224.
(40) Arey, J. S.; Martin Aparicio, A.; Vaiopoulou, E.; Forbes, S.; Lyon, D. Modeling the GCxGC Elution Patterns of a Hydrocarbon Structure Library To Innovate Environmental Risk Assessments of Petroleum Substances. *Environ. Sci. Technol.* **2022**, *56* (24), 17913−17923.
(41) Karolat, B.; Harynuk, J. Prediction of gas chromatographic retention time via an additive thermodynamic model. *J. Chromatogr. A* **2010**, *1217*, 4862−4867.
(42) McGinitie, T. M.; Harynuk, J. J. Prediction of retention times in comprehensive two-dimensional gas chromatography using thermodynamic models. *J. Chromatogr. A* **2012**, *1255*, 184−189.
(43) Zhang, X.; Ding, L.; Sun, Z.; Song, L.; Sun, T. Study on Quantitative Structure−Retention Relationships for Hydrocarbons in FCC Gasoline. *Chromatographia* **2009**, *70* (3−4), 511−518.
(44) Katritzky, A. R.; Chen, K.; Maran, U.; Carlson, D. A. QSPR correlation and predictions of GC retention indexes for methyl-branched hydrocarbons produced by insects. *Anal. Chem.* **2000**, *72* (1), 101−109.
(45) Yap, C. W. PaDEL-descriptor: an open source software to calculate molecular descriptors and fingerprints. *J. Comput. Chem.* **2011**, *32* (7), 1466−1474.
(46) *RDKit: Open-Source Chemoinformatics*, 2023. https://www.rdkit.org.
(47) Pearson, K. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. Lond. A* **1896**, *187*, 253−318.
(48) Spearman, C. The Proof and Measurement of Association between Two Things. *Am. J. Psychol.* **1904**, *15* (1), 72.
(49) Kendall, M. G. The treatment of ties in ranking problems. *Biometrika* **1945**, *33*, 239−251.
(50) Klein, C. T.; Kaiser, D.; Ecker, G. Topological distance based 3D descriptors for use in QSAR and diversity analysis. *J. Chem. Inf. Comput. Sci.* **2004**, *44* (1), 200−209.
(51) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(52) *scikit-learn 1.4.2: Machine Learning in Python*, 2024. https://scikit-learn.org/1.4/index.html.

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449

(53) Hoerl, A. E.; Kennard, R. W. Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics* **1970**, *12* (1), 69−82.
(54) Zschocke, A.; Scheuermann, S.; Ortner, J. *High Biofuel Blends in Aviation (HBBA)*, ENER/C2/2012/420-1 Final Report; Lufthansa & Wehrwissenschaftliches Institut für Werk- und Betriebsstoffe, 2012.

https://doi.org/10.1021/acs.energyfuels.4c05573
Energy Fuels 2025, 39, 4434−4449