

# Investigating the trace polar species present in diesel using high resolution mass spectrometry and selective ionization techniques

*Elize Smit* <sup>\*†</sup>, *Christopher P. Rüger* <sup>‡</sup>, *Martin Sklorz* <sup>‡§</sup>, *Stefan De Goede* <sup>||</sup>, *Ralf Zimmermann* <sup>‡§</sup>,  
*Egmont R. Rohwer* <sup>†</sup>

<sup>†</sup> Department of Chemistry, University of Pretoria, Lynnwood Road, Pretoria, South Africa

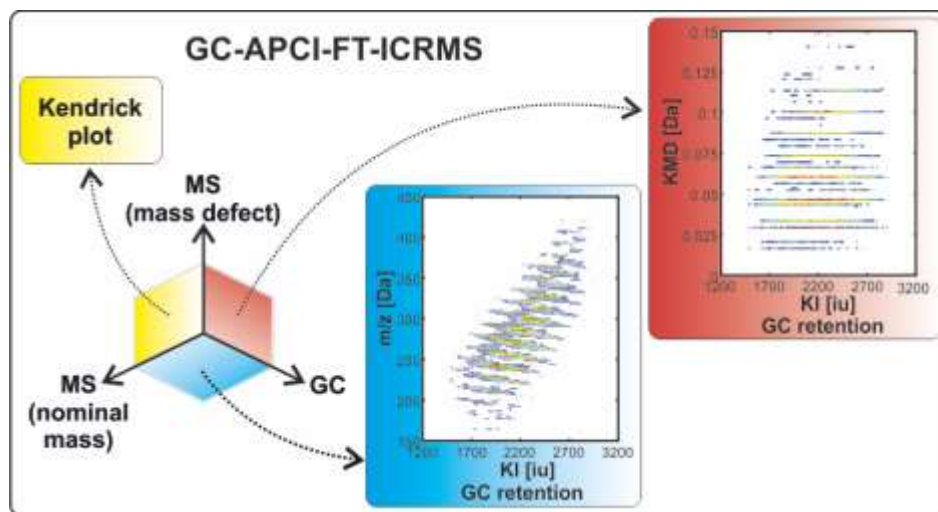
<sup>‡</sup> Joint Mass Spectrometry Centre / Chair of Analytical Chemistry, University of Rostock, Rostock, Germany

<sup>§</sup> Joint Mass Spectrometry Centre / Cooperation Group Comprehensive Molecular Analytics, Helmholtz Zentrum München, Neuherberg, Germany

<sup>||</sup> Sasol Southern Africa Energy, Energy Technology, Klasie Havenga Street, Sasolburg, South Africa

\* To whom correspondence should be addressed. E-mail: hanekom.elize@gmail.com. Telephone: +27762823911.

## Graphical Abstract



## **Abstract**

The trace polar species present in diesel were investigated by combining selective ionization with high resolution mass spectrometry. To eliminate matrix effects, the polar fraction was extracted using methanol and direct infusion with electrospray ionization and mass analysis were performed. The advantages and limitations of time-of-flight mass spectrometry and Fourier transform ion cyclotron resonance mass spectrometry for diesel analysis were discussed. Complementary information was obtained by considering both positive and negative ion mass spectra in terms of compounds that form part of the inherent fuel composition and compounds that represent fuel additives. Additionally, diluted fuels were separated by gas chromatography prior to atmospheric pressure chemical ionization and mass analysis. Results showed that a large portion of unsaturated polycyclic hydrocarbons were not detected during direct infusion experiments, indicating that these species were not extracted efficiently with methanol and/or ionized efficiently with electrospray.

## **Introduction**

Ultra high resolution mass spectrometry is often used for the analysis of crude oil and other petroleum derived fuels (e.g. heavy fuel oil) in petroleomics<sup>1-7</sup>. Fourier transform ion cyclotron resonance mass spectrometry (FT-ICRMS) is most commonly used, due to its high resolving power<sup>8-10</sup>. By employing selective ionization techniques, the in depth characterization of different species (or compound classes) of the crude oil is possible. For example, electrospray ionization (ESI) can be used to analyze polar species (e.g. acidic species readily form negative

ions)<sup>11,12</sup>, whilst atmospheric pressure chemical ionization (APCI) can be used to analyze less polarizable species, such as polycyclic aromatic species<sup>13</sup>. Non-polar species such as alkanes are not ionized during ESI, making this a selective ionization technique<sup>12</sup>.

ESI combined with quadrupole mass spectrometry has been used for the investigation of diesel samples with a focus on results from negative ion mass spectra<sup>14-16</sup>. Selective ionization combined with ultra-high resolution mass spectrometry is not often used for the analysis of middle distillate fractions such as diesel and only a limited number of examples can be found in literature<sup>17-19</sup>. This is mainly due to the fact that these fuels are highly suitable for gas chromatographic analysis considering its boiling point range (180 – 320 °C).

Comprehensive two dimensional gas chromatography (GC×GC) is ideal for the analysis of diesel, since high peak capacities (in terms of chromatographic separation) can be obtained<sup>20,21</sup>. Furthermore, GC×GC is often combined with electron impact ionization and time of flight mass spectrometry (TOFMS), which provides an additional dimension of chemical information. The bulk composition of diesel can easily be determined by GC×GC, since a characteristic pattern is observed in the resulting two-dimensional chromatogram<sup>20</sup>. It is also possible to characterize trace components present in diesel using GC×GC-TOFMS, although specialized data processing algorithms are necessary<sup>22,23</sup>. Detection of trace polar species may be troublesome and derivatization techniques<sup>24</sup> and/or selective detection<sup>25</sup> need to be employed.

Nevertheless, the characterization of trace polar species present in diesel is important, since these species contribute significantly to the physical properties of the fuel, including lubricity and stability<sup>26-28</sup>. This work shows how complementary information can be obtained from different analytical techniques with the objective to elucidate the composition of trace polar species in diesel. Methanol extracts of diesel samples were analyzed using ESI combined with

ultra-high resolution FT-ICRMS and high resolution TOFMS. Diesel samples were also analyzed by combining gas chromatographic separation with APCI and ultra-high resolution mass analysis (GC-APCI-FT-ICRMS), which has previously been used to study environmental samples<sup>29</sup>.

## Experimental Section

Four reference fuels were obtained from Haltermann GmbH (Germany) and Petrochem Carless (Belgium). These fuels were produced by blending refining streams in order to comply with the standard specifications of Europe (EN590), Sweden (Swedish MK1), USA (US 2-D) and California (California) respectively. Nine commercial diesels were bought from fuel stations in and around the Pretoria area in South Africa, i.e. from different suppliers in the same region. These commercial diesels will be referred to as commercial diesel X (CDX), where  $1 \leq X \leq 9$ . Two different diesel fuel grades were available at this time and are referred to as 50 ppm diesel or 500 ppm diesel (containing a maximum of 50 mg/kg or 500 mg/kg Sulphur respectively). All the commercial diesels were 50 ppm fuels except for CD8 and CD9 which were 500 ppm fuels. Solvents used were methanol (Romil-UpS<sup>TM</sup> Ultra Gradient for HPLC-MS, > 99.9 %), toluene (Sigma Aldrich Chromasolv<sup>®</sup> Plus for HPLC,  $\geq 99.9$  %) and acetonitrile (Romil-UpS<sup>TM</sup> Ultra Gradient for HPLC-MS, > 99.9 %).

The polar fraction of each diesel was extracted by mixing 1 ml of fuel with 2 ml of methanol. The methanol phase (top layer) was removed and diluted 1:3 (v/v) with acetonitrile for analysis by ESI-TOFMS. The methanol phase was diluted 1:10 (v/v) with methanol for analysis by ESI-FT-ICRMS. For analysis with GC-APCI-FT-ICRMS the neat fuel was diluted 1:10 (v/v) with a mixture of methanol and toluene (50:50), and this solution was further diluted 1:1 (v/v) with methanol.

**ESI-FT-ICRMS.** A Solarix FTICR MS system (Bruker Daltonics GmbH, Bremen, Germany) equipped with a 7 T magnet was used to analyze the samples. Bruker Compass Data Analysis 4.0 SP 5 software package (Bruker Daltonics GmbH, Bremen, Germany) was used to process acquired data. For the collection of positive ion mass spectra, samples were infused at a flow rate of 2.5  $\mu\text{l}/\text{min}$ . A voltage of +4 kV was applied between the capillary and inlet. The desolvation temperature was 300 °C. For the collection of negative ion mass spectra, samples were infused at a flow rate of 5  $\mu\text{l}/\text{min}$ . A voltage of -3 kV was applied between the capillary and inlet. The desolvation temperature was 200 °C. Collision induced dissociation (CID) was necessary to eliminate/minimize the formation of dimers. In source CID was set to 45 V, whilst quadrupole CID was set to 20 V. Mass to charge ratios ( $m/z$ ) between 100 and 1000 Da were recorded and 200 mass spectral scans were combined. Transient length per spectra was set to 0.490 s corresponding to one million data points (1M) and resulting in a resolution of roughly 150 000 @  $m/z$  200 FWHM (full width at half maximum). Total time of analysis was 2.5 minutes per sample.

An external calibration of the  $m/z$  scale was performed using arginine and its oligomeric clusters. Internal calibration was performed using a set of mass spectral peaks that were present in each of the samples and covered the mass range where peaks originating from diesel components were expected (Supplementary material, Table S1). Each of the samples was analyzed in triplicate.

Elemental compositions were assigned to mass spectral peaks with a signal-to-noise ratio (S/N) above six. Both odd and even electron configurations were allowed, with a double bond equivalent (DBE) range of -1.5 to 30 and mass accuracy was specified to be within 1 ppm. The following limits were applied to the range of atoms allowed: 0 – 50 C; 0 – 100 H; 0 – 2 N; 0 – 5

O, 0 – 5 S. For positive ion mass spectra, 0 – 1 Na atoms were also included to account for the formation of adducts.

**ESI-TOFMS.** A Synapt G2 HDMS system (Waters Inc., Milford, Massachusetts, USA) was used to analyze the samples. MassLynx™ (version 4.1) software (Waters Inc., Milford, Massachusetts, USA) was used to acquire and process data. The samples were infused at a flow rate of 5 µl/min. Capillary voltages of +4 kV and -2.5 kV were used to collect positive and negative ions respectively. The sampling cone was set to 40 V whilst the extraction cone was set to 5 V. The source temperature was 120 °C and the desolvation temperature was 300 °C. Mass to charge ratios ( $m/z$ ) between 50 and 1200 Da were recorded and 200 mass spectral scans were combined. An external calibration of the  $m/z$  scale was performed using sodium formate. Internal calibration was performed automatically by the software based on the measured and exact mass (lock mass) of leucine enkephalin ( $m/z$  555.2693), which was infused into the ion source throughout the analysis. Total time of analysis was 2.5 minutes per sample. All samples were analyzed in high resolution mode (longest flight path), resulting in resolution of 22 000 at  $m/z$  200 (FWHM).

Elemental compositions were assigned to mass spectral peaks with an absolute intensity above 3000. This arbitrary value was chosen based on close inspection of noise levels in the raw mass spectra. To predict elemental compositions the same limits were applied as for data obtained with ESI-FTICRMS, with the exception of mass accuracy, where a limit of 5 ppm was specified.

**GC-APCI-FT-ICRMS.** An Apex Qe Series II FTICR MS system (Bruker Daltonics, Billerica, MA, USA / Bruker Daltonics GmbH, Bremen, Germany) equipped with a 7 T magnet was used to analyze the samples. Bruker Compass Data Analysis 4.0 SP 5 software package (Bruker Daltonics GmbH, Bremen, Germany) was used to process acquired data. Separation was

performed using a CP 3800 gas chromatograph (Agilent, former Varian Technologies, Palo Alto, CA, USA) equipped with a PTV (programmed temperature vaporizing) injector. 1  $\mu$ l of the sample was placed into the ChromatoProbe device and inserted into the injector. The injector temperature was kept at 50 °C for 1 minute, heated at 10 °C/min to 80 °C and then heated at 60 °C/min to 320 °C, where it was kept for 20 minutes. A split ratio of 1:2 was utilized. Helium was used as carrier gas at a constant flow rate of 10 ml/min. A BPX5 column (SGE Analytical Science, Australia), 15 m  $\times$  250  $\mu$ m  $d_c$   $\times$  0.10  $\mu$ m  $d_f$ , was used. The column oven temperature was kept at 50 °C for 5 minutes, ramped at 5°C/min to 200 °C, then ramped at 10 °C/min to 250 °C, and finally ramped at 20 °C/min to 330 °C where it was kept for 8 minutes. The transfer line was kept at 320 °C. Total time of analysis was 52 minutes per sample.

The APCI source was adapted in order to ionize species in the gas phase<sup>30</sup>. The capillary voltage was -2 kV and a current of 3 $\mu$ A was applied to the corona needle. The APCI temperature was 320 °C. In source CID was set to 30 V in order to eliminate/minimize the formation of dimers/adducts. Mass to charge ratios ( $m/z$ ) between 120 and 1000 Da were recorded. Transient length per spectra was set to 1.15 s corresponding to two million data points (2M) resulting in a resolving power of roughly 340 000 @  $m/z$  200 and a scan frequency of 0.8 Hz. Internal calibration was performed using the same set of peaks used for data obtained with ESI-FT-ICRMS.

Elemental compositions were assigned to mass spectral peaks with a signal-to-noise ratio (S/N) above six. Both odd and even electron configurations were allowed, with a DBE range of -1.5 to 30 and mass accuracy specified to be within 1 ppm. The following limits were applied to the range of atoms allowed: 0 – 50 C; 0 – 100 H; 0 – 2 N; 0 – 5 O, 0 – 5 S, 0 – 5 Si. Initially the mass spectra collected throughout the GC run were combined and processed. Retention time

resolved data analysis was also applied in order to incorporate information obtained from the GC separation.

**Data processing.** The mass spectral data was processed based on a procedure described in literature<sup>31</sup>. To our knowledge this approach has not been applied very often in literature, but it is especially useful when data is obtained from instruments with limited resolving power. Kendrick mass defects (KMDs)<sup>32</sup> and nominal mass series ( $z^*$ ) values<sup>31</sup> were calculated using the measured  $m/z$  values. Peaks were assigned to the same subclass (homologous series) if their  $z^*$  values were identical (i.e. their nominal masses differed by multiples of 14), and their KMDs differed less than 1 mDa. The predicted elemental compositions were used to obtain a general formula, where  $C_nH_{2n+z}X_x$  is abbreviated to  $ZX_x$ <sup>11,33</sup>. For example, the quasi-molecular ions  $[M-H]^+$  of palmitic and stearic acid, would have assigned elemental compositions of  $C_{16}H_{31}O_2$  and  $C_{18}H_{35}O_2$  respectively, but both would have the same general formula, i.e.  $-1O_2$ . A general formula was assigned to each subclass (homologous series) based on the most frequent general formula assigned to each of the peaks thereof. By doing this, incorrect assignments (especially at higher masses) could be eliminated, to a certain extent, in cases where limited resolving power was available. Furthermore, elemental compositions could be assigned to some species that were originally unassigned (given that the new assignment had a mass accuracy of  $\pm 5$  ppm). The intensities of species were normalized to total intensity of peaks for which elemental compositions were assigned. Note that peaks representing compounds containing  $^{13}C$  isotopes were removed before normalization.

The relative intensities of all the mass spectral peaks (quasi-molecular ions) in a homologous series were summed to get the composition of each sample in terms of subclasses of ions. This information was then summarized further by considering only the heteroatom content to obtain

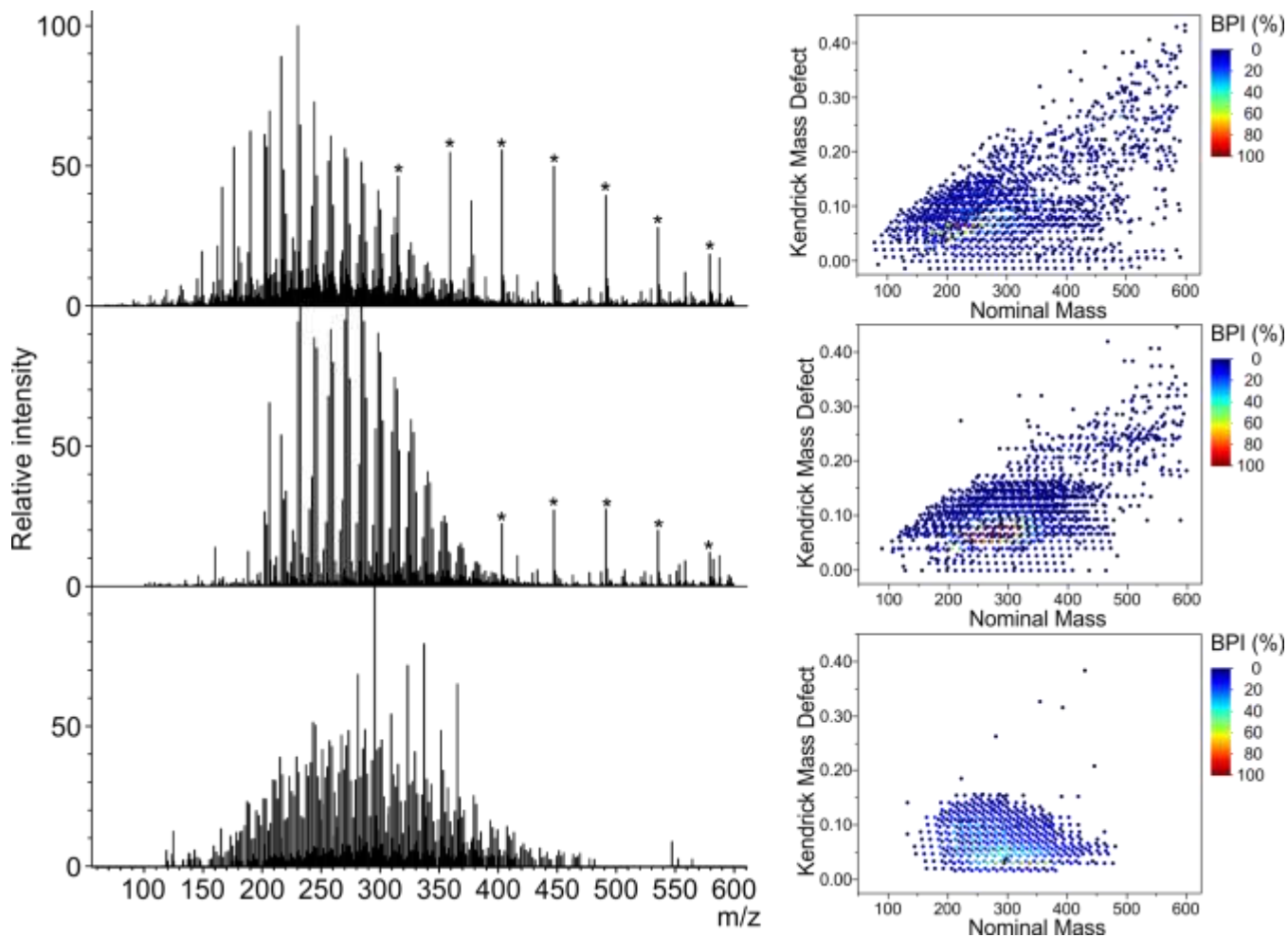


the composition in terms of classes of ions, e.g. CH and CHO<sub>x</sub> (where 1 ≤ X ≤ 5). For samples analyzed in triplicate, the average sample composition was determined by considering only species (based on assigned formulae) that were present in at least two of the three replicates, therefore eliminating noise and/or incorrectly assigned peaks.

Time-resolved data analysis was performed for the GC-APCI-FT-ICRMS data. Automatic feature detection, i.e. region of interest at the *m/z* and retention time scale, was done by applying a MATLAB algorithm (developed in-house) based on the approach of Tautenhahn et al.<sup>34</sup> using a *m/z*-window of 5 ppm. Briefly, the algorithm detects mass traces using a moving window where the center is calculated as the average position of the peaks. This approach has the benefit that no fixed binning size needs to be specified. Chromatographic features of unresolved time traces were calculated using simple smoothing and local maximum detection. Relevant features were obtained by applying an intensity threshold, removing of isotopes and ionization-artifacts (O-adducts) and removing known Si-containing species originating from column bleed. Spectra number was calculated into time domain and consequently converted into Kováts retention index (KI) using an external calibration with a PAH-mixture and the relation between Lee-Index and Kováts-Index<sup>35,36</sup>.

## Results and Discussion

**Comparison of analytical techniques.** The positive ion mass spectra obtained for commercial diesel 5 (CD5) from the different techniques are shown in Figure 1. CD5 was selected to illustrate results since it represents the sample set well. The mass spectrum in Figure 1c was obtained by averaging the mass spectral scans collected throughout the GC run. Different distributions were obtained, which can be attributed to several factors. The geometry of the



**Figure 1.** Positive ion mass spectra (left) and Kendrick plots (right) obtained for CD5 analyzed with ESI-TOFMS (top), ESI-FT-ICRMS (middle) and GC-APCI-FT-ICRMS (bottom) to illustrate the change in ion profile of various techniques. The peaks marked with \* represent polyethylene glycol, a common contaminant in ESI.

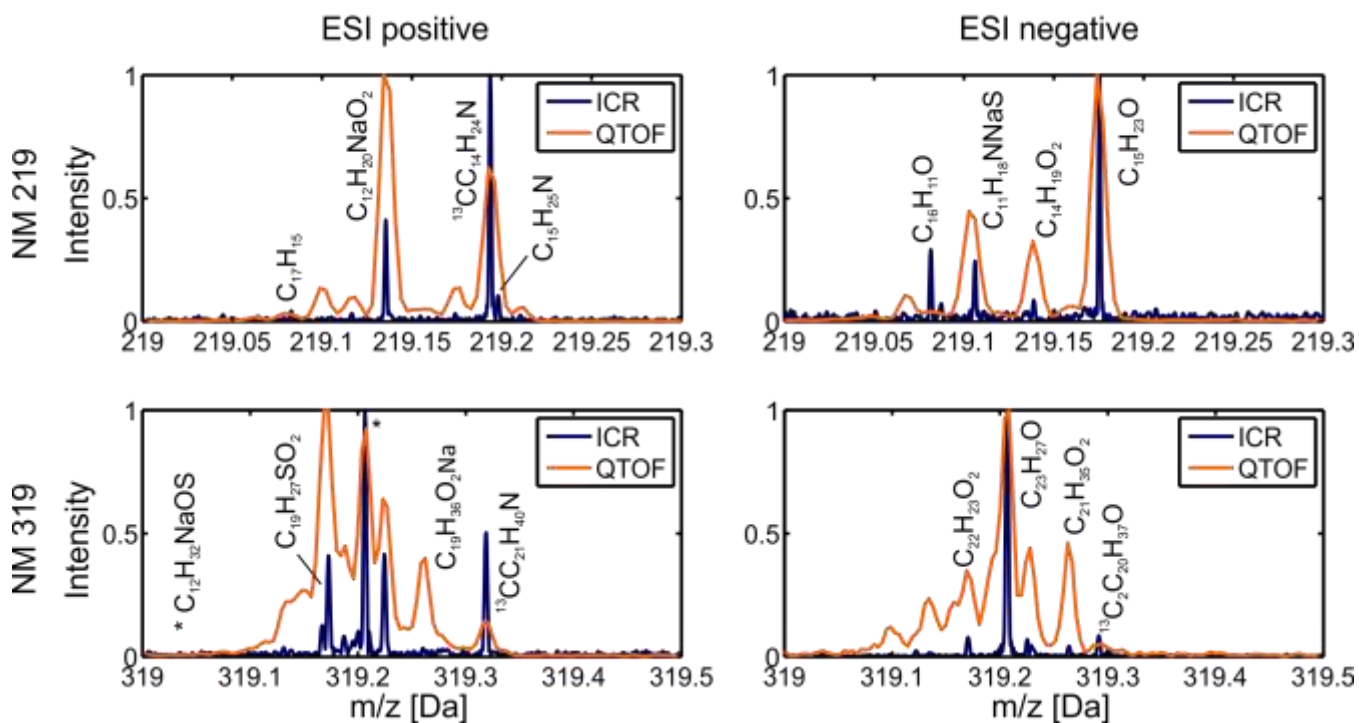
different ESI sources may influence ion transmission efficiency. Ions across the whole mass range are transmitted to the detector with the same efficiency when using ESI-TOFMS (Figure 1a), whilst the transmission of lower mass ions into the ICR cell is limited for ESI-FT-ICRMS (Figure 1b). Specific parameters of the FT-ICRMS (e.g. the delay between ion accumulation and trapping) may be tuned to shift the distribution towards lower masses, but this could result in the limited transmission (or exclusion) of higher mass ions. Specific settings used during FT-ICRMS experiments are given in the supporting material (Table S2). It is hardly feasible to transmit ions

below 100 Da efficiently into the ICR cell by tuning instrument parameters. Some peaks with higher  $m/z$  values differ by 44.0263 Da, which corresponds to  $C_2H_4O$ , indicating that these peaks may represent polyethylene glycol (with an increasing number of monomers) which is a common contaminant in ESI<sup>37,38</sup>.

Different Kendrick plots were obtained from ESI and APCI experiments (Figure 1) since different species tend to ionize with these techniques, and these results will be discussed in detail below. Interestingly, the Kendrick plots observed for the two ESI-MS experiments were similar. This could indicate that comparable mass accuracies were obtained with the different analyzers despite the significant difference in their resolving power. Closer inspection of these plots reveal that less variation in the KMD-values (for a specific homologous row) was observed with FT-ICRMS, especially at higher masses (i.e.  $m/z > 350$ ). To investigate this further, the mass resolution and mass accuracy of the analyses with ESI-TOFMS and ESI-FT-ICRMS were compared by investigating the raw mass spectra at low and high  $m/z$  values in Figure 2. Even though the resolving power of TOFMS is not comparable to that of FT-ICRMS,  $m/z$  values could still be measured accurately given that the sample complexity is limited. Here, the distinction between mass accuracy and mass resolution is important, i.e. if a mixture is not too complex and mass spectral peaks do not overlap, the  $m/z$  ratio can be measured accurately. The accuracy of determining the centroid of a peak is not influenced by the peak width, although less peak overlap will be observed for narrower peaks (this is evident for higher masses).

At lower  $m/z$  values, similar species were observed with both analyzers, and for the  $m/z$  window shown in Figure 2 (top) only one peak ( $C_{15}H_{25}N$ ) was not resolved by ESI-TOFMS. At higher  $m/z$  values the resolving power of the FT-ICRMS analyzer is crucial for distinguishing between different species and consequently determining their  $m/z$  values accurately, as observed

in the Kendrick plots (Figure 1). The sophisticated peak picking algorithm of the MassLynx<sup>TM</sup> software, the use of an internal calibrant and data processing based on KMD and nominal mass series aids in the detection and accurate assignment of some partially overlapping peaks at a nominal mass of 319 Da (see Figure 2). Complementary information can thus be obtained: it is possible to determine the true  $m/z$  distribution (without transmission bias) of the trace polar species using high resolution ESI-TOFMS, whilst the ultra-high resolution obtained with ESI-FT-ICRMS provides important information at especially higher masses.



**Figure 2.** Comparison of mass resolution and mass accuracy obtained at low (top) and high (bottom) masses using ICRMS and TOFMS. The positive (left) and negative (right) ion mass spectra of CD5 are shown.

The information obtained from the different analytical techniques can further be compared by considering several values describing the  $m/z$  distribution as well as chemical information, as shown in Table 1. Values were determined after data processing methods (used for the assignment of peaks to subclasses and classes) were applied to the respective data sets (see

Experimental section). The base peak, mean  $m/z$  and number of C-atoms (weighted by intensity) confirm that the ICR transmission is biased towards higher masses. The maximum number of peaks detected per nominal mass (rounded Kendrick masses were used to determine nominal masses) gives an indication of sample complexity, and more chemical species were observed with ESI (positive ions) compared to ESI (negative ions) and APCI (positive ions). It is interesting to note that the same maximum number of peaks per nominal mass were obtained for ESI-TOFMS and ESI-FT-ICRMS although the  $m/z$  windows differ, which corresponds to previous discussions regarding the transmission of ions and resolving power. From these results it is also clear that the complexity of diesel is not comparable to that of crude oil or heavier distillates since up to 15 peaks per nominal mass have been detected in crude oil<sup>33</sup> and asphaltene fractions<sup>11</sup>. The results showed that high resolution TOFMS may be sufficient for the analysis of diesel samples, but this may not be the case for more complex samples such as crude oil where higher resolving power is needed to resolve the multiplets at high  $m/z$  values.

The mean DBE values shows that positive ions detected with ESI-MS were more unsaturated than negative ions detected with ESI-MS and positive ions detected with APCI-MS. This correlates with the mean hydrogen to carbon ratios (H/C). Acidic O-containing species readily form negative ions during electrospray, and especially the ESI-TOFMS results seem to indicate that these species are more saturated than the basic N-containing species that readily form positive ions during electrospray (see Table 1). For the ESI-FT-ICRMS results the discrepancy between the mean DBE values for positive and negative spectra was less pronounced, but inspection of data shown later (Figure 5) reveals that the  $\text{CHN}_1$  class contains less saturated species than the  $\text{CHO}_2$  class. Species containing no heteroatoms were more readily ionized with

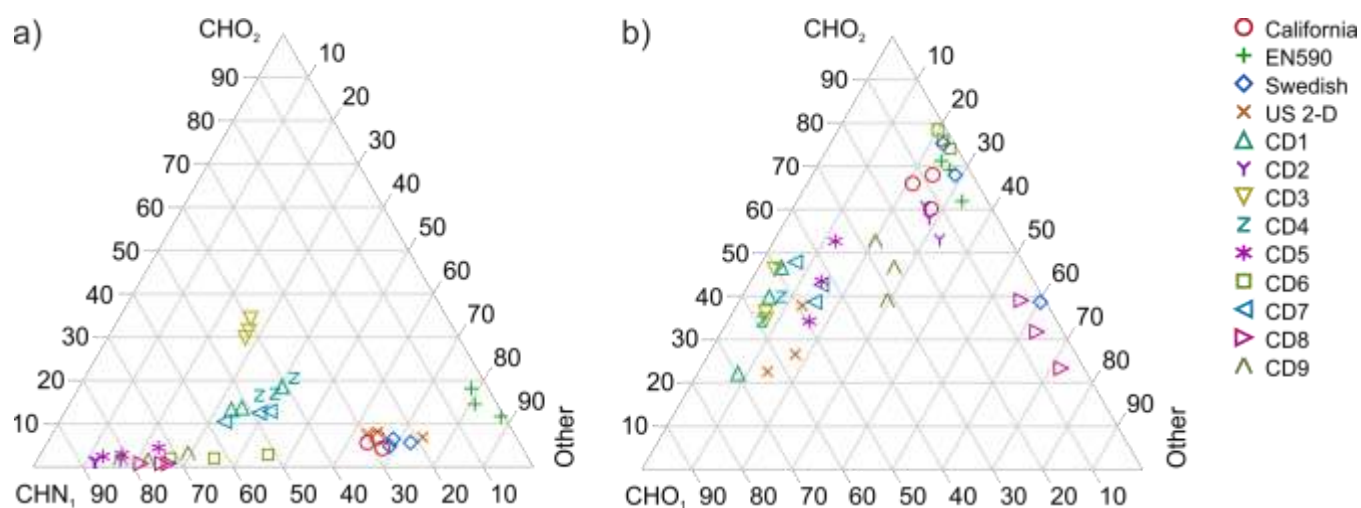
**Table 1.** Comparison of CD5 composition results obtained using different analytical techniques.

Instrument	ESI-TOFMS		ESI-FTICRMS		GC-APCI-FTICRMS
Ions	Positive	Negative	Positive	Negative	Positive
Sample preparation	Methanol extract	Methanol extract	Methanol extract	Methanol extract	Diluted fuel
Base peak	230.1903	281.2486	284.2373	281.2486	295.2996
	C <sub>16</sub> H <sub>24</sub> N	C <sub>18</sub> H <sub>33</sub> O <sub>2</sub>	C <sub>20</sub> H <sub>30</sub> N	C <sub>18</sub> H <sub>33</sub> O <sub>2</sub>	C <sub>20</sub> H <sub>39</sub> O
Mean $m/z^a$	278	295	301	313	290
Mean nr of C-atoms <sup>a</sup>	18	18	21	21	20
Maximum nr of peaks per nominal mass (processed data)	7 @ $m/z$ 255, 280, and 285	6 @ $m/z$ 207 and 225	7 @ $m/z$ 311, 321 and 349	6 @ $m/z$ 353	6 @ $m/z$ 259 and 297
Major class and percentage of total ion intensity	CHN <sub>1</sub> 58 %	CHO <sub>2</sub> 61 %	CHN <sub>1</sub> 82 %	CHO <sub>2</sub> 47 %	CH 53 %
Mean DBE <sup>a</sup>	5.46	3.87	6.02	5.43	4.79
Mean H/C <sup>a</sup>	1.553	1.695	1.548	1.585	1.607
Mean O/C <sup>a</sup>	0.044	0.100	0.020	0.076	0.034
Mean N/C <sup>a</sup>	0.043	0.012	0.044	0.005	0.0001
Total number of peaks observed (processed data)	1324	517	919	577	479

<sup>a</sup> weighted by intensity

APCI as expected and these species make up 53 % of the total ion intensity. Besides the fact that APCI has a high selectivity for polycyclic hydrocarbons, the results also reflect that the sample was only diluted and methanol extraction of polar species was not performed prior to analysis (see Supporting material, Figure S1). The different species observed with ESI and APCI also explains the discrepancy between the Kendrick plots in Figure 1.

**ESI-FT-ICRMS results.** Each of the samples was analyzed in triplicate using ESI-FT-ICRMS. Ternary plots may be used to represent the components of a mixture on three axes, limiting all the possible combinations of these components to the space within a triangle. The position of a point in the triangle represents the composition of a mixture, and therefore points that are close to one another represent mixtures with similar compositions. Ternary plots were used to illustrate the sample composition in terms of classes and to investigate the repeatability (Figure 3), since the ion intensities were normalized in terms of total composition (i.e. total

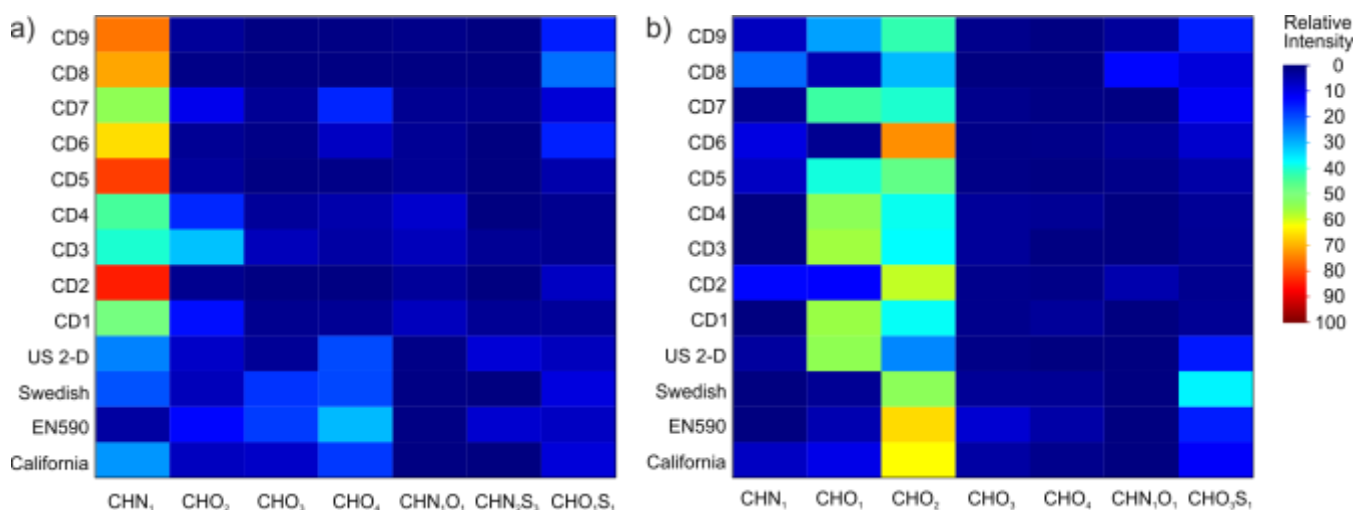


**Figure 3.** Ternary plots showing the variation of replicates in terms of classes obtained from positive (a) and negative (b) ion mass spectra.

intensity of observed ions) and therefore the variables (relative intensity of classes) are dependent on one another, making the interpretation of standard deviations difficult. The major classes in the positive ion mass spectra were  $\text{CHN}_1$  and  $\text{CHO}_2$ , whilst the major classes in the negative ion mass spectra were  $\text{CHO}_2$  and  $\text{CHO}_1$ . The remaining classes were summed together and used as the third variable (axis) in the ternary plot.

The reference diesels contained relatively smaller amounts of  $\text{CHN}_1$  species compared to the commercial diesels in the positive ion mass spectra. Similar sample compositions were obtained for the replicates of most samples, especially for the positive ion mass spectra. The composition of Swedish diesel varied a lot for the negative ion mass spectra, especially in terms of  $\text{CHO}_1$  and the remaining classes. This is reasonable considering that the Swedish diesel contained much less aromatic and polar/polarizable species compared to the other samples, resulting in very low ion intensities.

In order to investigate the sample compositions in more detail, the average relative intensities ( $n = 3$ ) of the most abundant classes were determined for the different samples and are shown in Figure 4. Although only the major ion classes are shown in Figure 4 and discussed here, other ion classes (e.g.  $\text{CHN}_2$ ) were also observed during analyses. Each of the classes represented here was selected based on the fact that they were one of the three most abundant classes in at least one of the samples.



**Figure 4.** Comparison of relative intensities of the major ion classes obtained for the different reference and commercial diesels for positive (a) and negative (b) ions detected by ESI-FT-ICRMS.

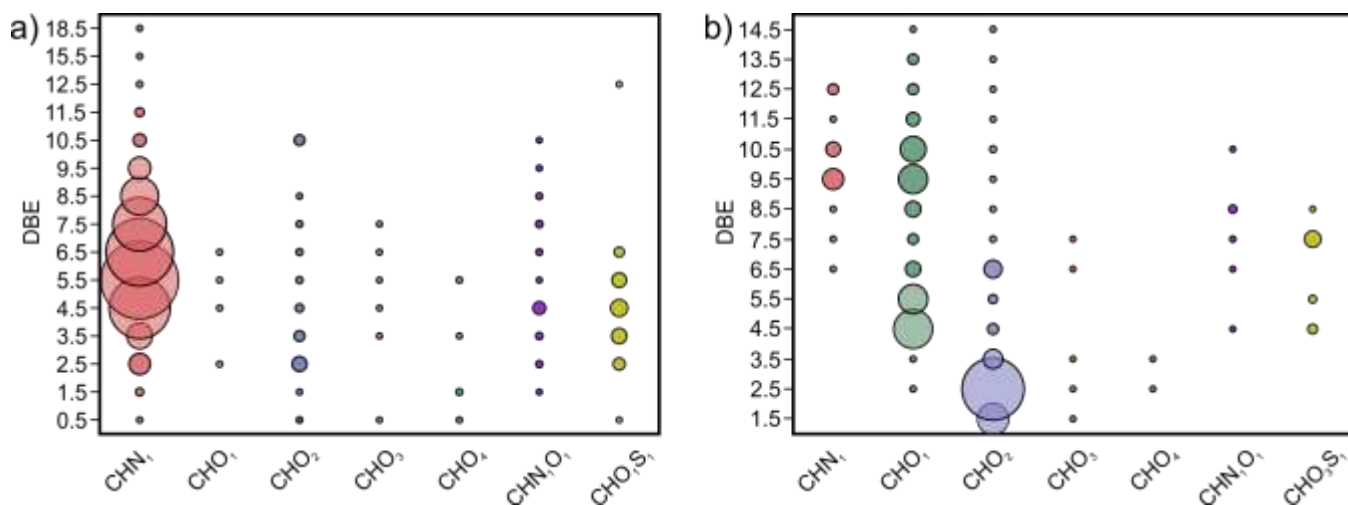


The positive ion mass spectra (Figure 4a) of the commercial diesels display mostly ions of compounds containing only one N-atom ( $\text{CHN}_1$  class). For the reference diesels, other classes such as  $\text{CHO}_4$  were present in relatively larger amounts as well. These chemical differences can be explained by considering the origins of reference and commercial fuels. Reference diesels are produced by blending certain refinery streams to obtain desired physical properties based on standard specifications, whereas commercial diesels are distillation fractions of crude oil (or other fossil fuel derived products) which contain additives that ensure the physical properties of the fuel comply with standard specifications. Other classes such as  $\text{CHO}_2$  and  $\text{CHO}_1\text{S}_1$  also contributed significantly to the chemical composition of most samples. The  $\text{CHO}_1\text{S}_1$  class may represent species formed by the oxidation of S-containing species during storage<sup>18,39</sup>.

Furthermore the negative ion mass spectra (Figure 4b) display mostly ions of compounds containing two O-atoms ( $\text{CHO}_2$  class). In this case the differences between commercial and reference diesels are more subtle. Other classes that contribute to the chemical composition of the samples are  $\text{CHO}_1$  and  $\text{CHO}_3\text{S}_1$ . The  $\text{CHO}_3\text{S}_1$  class is relatively more prominent in the reference diesels than most of the commercial diesels. The  $\text{CHO}_1$  class was not determined to be one of the major classes for the positive ion mass spectra based on the criteria specified above.

Complementary information is obtained by considering both positive and negative ion mass spectra. Trace polar species that influence the physical properties of a fuel due to its basic or acidic properties will preferentially form positive or negative ions respectively. The most abundant class in the negative ion mass spectra of most samples were  $\text{CHO}_2$ , and consist mainly of a base peak with  $m/z$  value of either 281.248 or 279.232 which represent quasi-molecular ions,  $[\text{M}-\text{H}]^-$ , of  $\text{C}_{18}\text{H}_{32}\text{O}_2$  and  $\text{C}_{18}\text{H}_{34}\text{O}_2$  respectively. These species may represent long chain carboxylic acids (or esters) that are commonly used as lubricity additives in diesel<sup>27</sup>.

The sample composition can be investigated further by considering the DBE distribution for each class. The results obtained for CD5 are shown in Figure 5 and represents the mean values observed for the three replicates. For the positive ion mass spectrum (Figure 5a), the  $\text{CHN}_1$  class has DBE values ranging from 0.5 to 18.5, which represents quasi-molecular ions of compounds containing between 1 and 19 double bonds and/or rings. The DBE distribution of this class follows a pattern that can be expected for a mixture obtained by isolating a specific distillation fraction, indicating that these species form part of the inherent fuel composition and do not represent additives that were added to comply with fuel specifications. The other classes were not nearly as abundant as the  $\text{CHN}_1$  class. The  $\text{CHN}_1$  class was preferentially ionized by ESI+ and therefore simultaneously may have suppressed ionization of other classes. The distributions of the other classes also suggest that they most likely form part of the inherent fuel composition.



**Figure 5.** DBE distribution of the major classes for CD5 for both positive (a) and negative (b) ions detected by ESI-FT-ICRMS. Each of the circles represents the total ion intensity of a homologous series (KMD row), whilst the size of each circle represents its relative abundance.

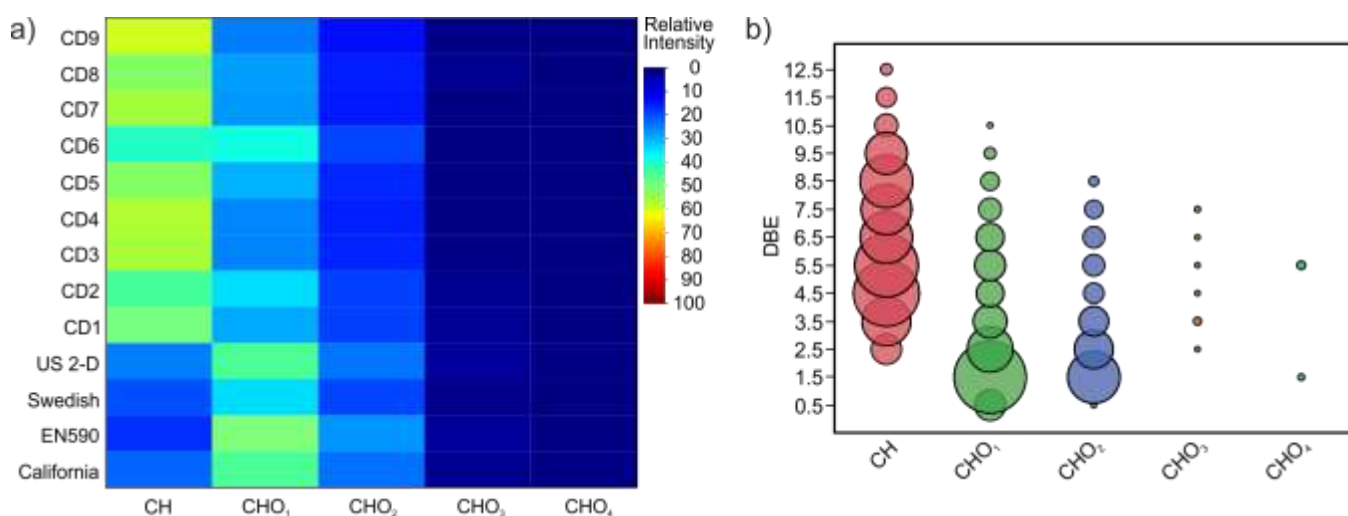
Closer inspection of the DBE distribution of classes observed in the negative ion spectra (Figure 5b) reveals that the most abundant subclass represents compounds containing two O-atoms and two double bonds and/or rings. The most abundant subclasses in the CHO<sub>1</sub> and CHN<sub>1</sub> classes seem to be highly unsaturated compared to the most abundant subclasses in the CHO<sub>2</sub> class. The non-uniform DBE distributions of the ion classes may indicate that certain additives preferentially form negative ions during electrospray, as discussed for the CHO<sub>2</sub> class above. The most abundant subclasses in the CHO<sub>3</sub>S<sub>1</sub> class have DBE values of 7.5 and 4.5, which may represent alkylated naphthalene sulfonate and benzene sulfonate compounds respectively. Even though these compounds are not commonly used as diesel additives, they play an important role in oil recovery processes<sup>40</sup>.

It is interesting to note that both CHN<sub>1</sub> and CHO<sub>2</sub> classes (among others) were observed in both positive and negative ion mass spectra, but their distributions are not the same. This may indicate that these species are amphiprotic to a certain extent (if similar DBE distributions were observed). In cases where the DBE distributions differ significantly (for positive and negative ion mass spectra) it is possible that compounds with different functional groups were detected. Structural differences may cause compounds with the same general formula to form either positive or negative ions, e.g. carboxylic acids will readily form negative ions whilst esters will more likely form positive ions even though both species belong to the CHO<sub>2</sub> class.

**GC-APCI-FT-ICRMS results.** Initially the mass spectra collected throughout each of the GC-APCI-FT-ICRMS analyses were combined to obtain one mass spectrum for each sample. The major class detected for the reference diesels consisted of hydrocarbons without any heteroatoms (CH class), whilst the CHO<sub>1</sub> class was relatively more abundant for the reference diesels (Figure 6a). This analytical technique provides information that is complementary to the

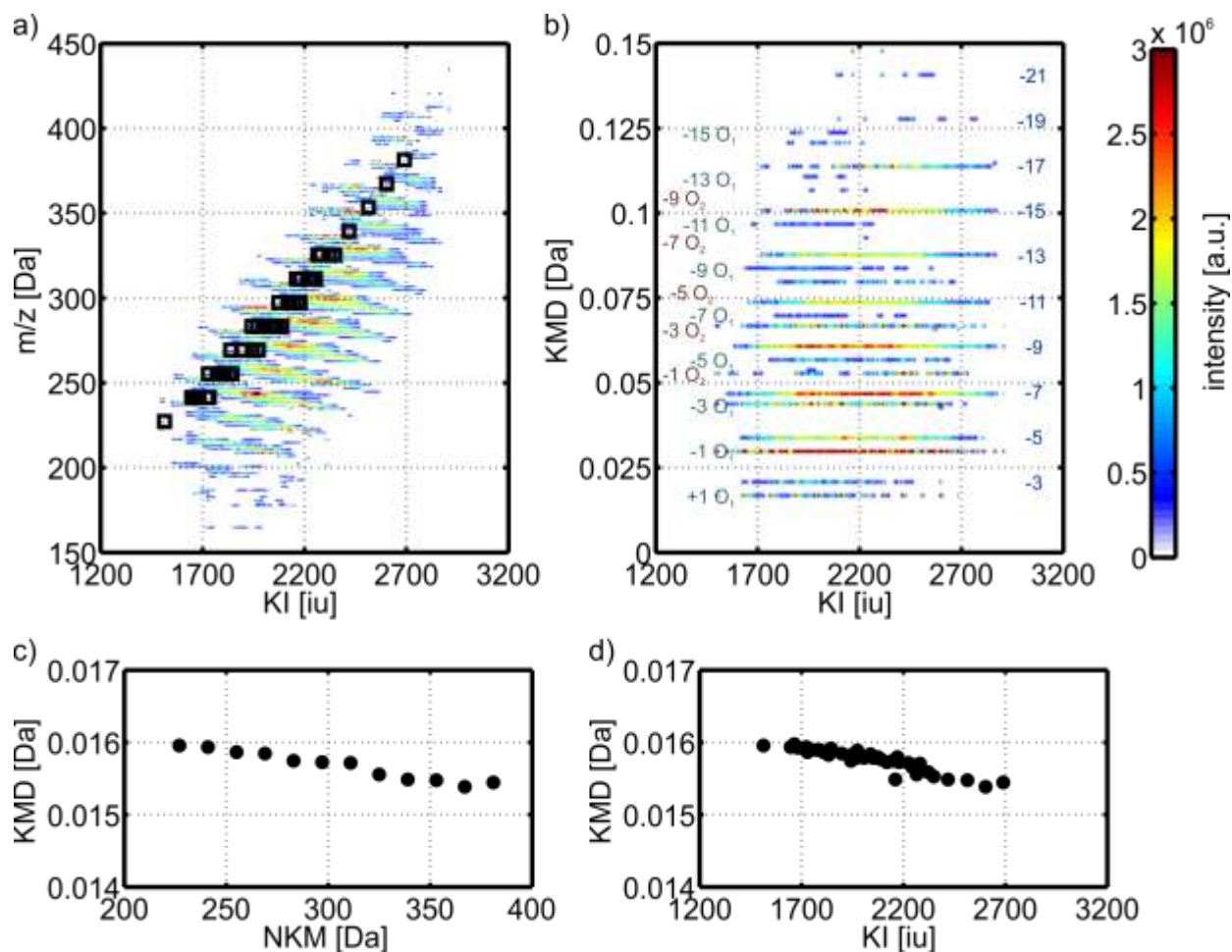
results obtained from ESI-FT-ICRMS and ESI-TOFMS. Separation by GC minimizes matrix effects which may cause ionization suppression and therefore methanol extraction was not necessary. Furthermore, other classes were observed in the positive ion mass spectra as a result of the different ionization method. This may also be due to the fact that unsaturated hydrocarbon species (CH class) may not be extracted as efficiently as other classes with methanol (see Supporting material, Figure S1).

The DBE distributions of the classes of CD5 are shown in Figure 6b. The results indicate that these species form part of the inherent fuel composition, based on the pattern of the DBE distribution as discussed before. Based on the DBE values, the CH class consists of species with 3 to 13 double bonds and/or rings, therefore mostly representing alkylated rows of aromatic species. The CHO<sub>1</sub> and CHO<sub>2</sub> classes contain species that are slightly more saturated. These classes may represent species with several different functional groups and therefore structural information is necessary, which was disregarded by combining the mass spectra acquired throughout each analysis.



**Figure 6.** (a) Comparison of the major classes obtained for the different reference and commercial diesels using GC-APCI-FT-ICRMS. (b) DBE distribution of the major classes for CD5.

Retention time resolved data analysis was performed on the GC-APCI-FT-ICRMS data (similar to procedures used in conventional GC-MS) in order to incorporate the structural information obtained by gas chromatographic pre-separation. Figure 7 shows the results obtained for CD5, where each spot represents a deconvoluted mass spectral peak (features detected during data processing).



**Figure 7.** Different graphical representations of the chromatographic separation and mass spectral data obtained by GC-APCI-FT-ICRMS for sample CD5. (a)  $m/z$  vs. KI values, with peaks of the  $+1O_1$  subclass highlighted in black. (b) Plot of Kendrick mass defect vs. KI values showing a more orthogonal distribution of peaks throughout the two-dimensional space. Comparison of the  $+1O_1$  subclass with (d) and without (c) chromatographic separation taken into account.

Peaks with the same accurate  $m/z$  values have different KI values (Figure 7a), therefore indicating that isomers and/or compounds with different functional groups were separated to a certain extent. Based on the non-polar stationary phase used for the chromatographic separation, the more retained species (higher KI values) have lower vapor pressures. As an example, the +1O<sub>1</sub> subclass is highlighted in black, where each spot represents a deconvoluted mass spectral peak. This subclass (C<sub>n</sub>H<sub>2n+1</sub>O) represents quasi-molecular ions of species with the general formula C<sub>n</sub>H<sub>2n</sub>O, i.e. species containing one O-atom and one double bond or ring, e.g. aldehydes, ketones or alkenes with a hydroxyl or ether group. For this specific subclass, at least seven isomers were detected for a single  $m/z$  value.

The multidimensional data obtained from GC-APCI-FT-ICRMS may be compared with GC×GC-TOFMS where electron ionization (EI) is routinely employed. For the analysis of diesel with GC×GC-TOFMS, a non-polar stationary phase is typically used in the first dimension together with a polar stationary phase in the second dimension<sup>21</sup>. Information regarding vapor pressure (retention time in the first dimension), polarity (retention time in the second dimension),  $m/z$  (nominal mass resolution) and relative intensities can thus be obtained. Similar information can be obtained when using GC-APCI-FT-ICRMS for the analysis of diesel, although the selectivity of the ionization method should be taken into account. The vapor pressures of species are reflected in the retention time (or KI values) and relative intensities are obtained. Accurate mass measurement provides information regarding the number of nucleons in the molecule (nominal mass) together with information regarding polarity in terms of heteroatom composition (based on nuclear binding energy). In GC×GC-TOFMS it is important that the mechanisms of separation in the two chromatographic dimensions differ in order to obtain an orthogonal system, resulting in a high peak capacity. The orthogonality of the system can be quantified by

considering the degree of linear association (by calculating the Pearson's correlation coefficient) between the two dimensions<sup>41</sup>. In Figure 7a a linear association exists between the  $m/z$  and KI values, since the mechanism of separation was based on vapor pressure. This means that the system is not orthogonal.

By considering the KMD values for each of the species against retention time (Figure 7b), the peaks are distributed throughout the two-dimensional space, resulting in an orthogonal system. Subclasses are grouped together in a single horizontal line and the resulting plot resembles a conventional Kendrick plot where the x-axis is based on nominal mass rather than vapor pressure (Figure 1). The main advantage is that structural information obtained from the GC separation is incorporated in the x-axis. Some of the major classes (and their respective subclasses) have been labelled. The species that were highlighted in Figure 7a are now present in a single horizontal line. The most abundant subclass is  $-1O_1$  (with a DBE value of 1.5) which corresponds to the results shown for the combined mass spectral results (Figure 6b).

Figures 7 c and d show the  $+1O_1$  subclass where KMD values are plotted against nominal mass and KI values respectively, showing in detail the large amount of information obtained by doing retention time resolved data analysis. This specific subclass consists of 13 peaks when the GC separation is not taken into account and 42 peaks when retention time resolved data analysis is performed. These results also show that GC pre-separation increases the amount of information obtained, even though much longer analysis times are required.

## Conclusions

This work shows that it is possible to analyze the trace polar species present in diesel by employing selective ionization techniques combined with high resolution mass spectrometry.

CHN<sub>1</sub> and CHO<sub>2</sub> species tend to form positive and negative ions respectively during ESI-MS, whilst CH species were predominantly observed during APCI-MS.

Complementary information was obtained from two different mass analyzers. It is suggested that high resolution TOFMS may be sufficient for the analysis of middle distillates and may be especially useful for the characterization of lighter distillation fractions such as gasoline, although it will not be suitable for highly complex samples such as crude oil which contain higher mass compounds. The main advantage of TOFMS is that the ion transmission is unbiased, whilst the main advantage of FT-ICRMS is its unsurpassed resolving power.

Results indicate that species from the inherent fuel composition tend to form positive ions (during both ESI and APCI), whilst additives tend to form negative ions. Furthermore, the positive ion mass spectra tend to be more complex compared to the negative ion mass spectra. The combination of GC with APCI-FT-ICRMS provides information regarding the structural features of isobaric compounds. This multidimensional information can be visualized by using a variation of the traditional Kendrick plot, where GC retention times are used instead of nominal mass values on the x-axis.

**Acknowledgements.** The financial assistance of Sasol (Sasol Southern Africa Energy, Energy Technology), the Deutscher Akademischer Austausch Dienst (DAAD) and the National Research Foundation (DAAD-NRF) towards this research is hereby acknowledged as well as the German Science Foundation (DFG) for funding the Bruker Solarix FTICR MS. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the DAAD-NRF.



## References

- (1) Marshall, A. G.; Rodgers, R. P. *Acc. Chem. Res.* **2004**, *37* (1), 53–59.
- (2) Marshall, A. G.; Rodgers, R. P. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (47), 18090–18095.
- (3) Rodgers, R. P.; Schaub, T. M.; Marshall, A. G. *Anal. Chem.* **2005**, *77* (1), 20 A – 27 A.
- (4) Panda, S. K.; Andersson, J. T.; Schrader, W. *Anal. Bioanal. Chem.* **2007**, *389* (5), 1329–1339.
- (5) Bae, E.; Na, J.-G.; Chung, S. H.; Kim, H. S.; Kim, S. *Energy & Fuels* **2010**, *24* (4), 2563–2569.
- (6) Fasciotti, M.; Lalli, P. M.; Klitzke, C. F.; Corilo, Y. E.; Pudenzi, M. A.; Pereira, R. C. L.; Bastos, W.; Daroda, R. J.; Eberlin, M. N. *Energy and Fuels* **2013**, *27* (12), 7277–7286.
- (7) Klitzke, C. F.; Corilo, Y. E.; Siek, K.; Binkley, J.; Patrick, J.; Eberlin, M. N. *Energy Fuels* **2012**, *26* (9), 5787–5794.
- (8) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.
- (9) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25* (2), 282–283.
- (10) Cho, Y.; Ahmed, A.; Islam, A.; Kim, S. *Mass Spectrom. Rev.* **2015**, *34* (2), 248–263.
- (11) Qian, K.; Robbins, W. K.; Hughey, C. A.; Cooper, H. J.; Rodgers, R. P.; Marshall, A. G. *Energy and Fuels* **2001**, *15* (6), 1505–1511.
- (12) Zhan, D.; Fenn, J. B. *Int. J. Mass Spectrom.* **2000**, *194* (2-3), 197–208.
- (13) Herrera, L. C.; Grossert, J. S.; Ramaley, L. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1926–1941.
- (14) Rostad, C. E. *Energy and Fuels* **2005**, *19* (3), 992–997.
- (15) Rostad, C. E. *Environ. Forensics* **2006**, *7* (1), 5–14.
- (16) Rostad, C. E.; Hostettler, F. D. *Environ. Forensics* **2007**, *8* (1-2), 129–137.
- (17) Haddad, R.; Regiani, T.; Klitzke, C. F.; Sanvido, G. B.; Corilo, Y. E.; Augusti, D. V.; Pasa, V. M. D.; Pereira, R. C. C.; Romão, W.; Vaz, B. G.; Augusti, R.; Eberlin, M. N. *Energy and Fuels* **2012**, *26* (6), 3542–3547.
- (18) Liu, P.; Xu, C.; Shi, Q.; Pan, N.; Zhang, Y.; Zhao, S.; Chung, K. H. *Anal. Chem.* **2010**, *82* (15), 6601–6606.
- (19) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G. *Energy and Fuels* **2001**, *15* (5), 1186–1193.
- (20) Phillips, J. B.; Xu, J. *J. Chromatogr. A* **1995**, *703* (1-2), 327–334.
- (21) Dallüge, J.; Beens, J.; Brinkman, U. A. T. *A Century Chromatogr. 1903-2003* **2003**, *1000* (1–2), 69–108.

- (22) Welthagen, W.; Schnelle-Kreis, J.; Zimmermann, R. *J. Chromatogr. A* **2003**, *1019* (1-2), 233–249.
- (23) Striebich, R. C.; Contreras, J.; Balster, L. M.; West, Z.; Shafer, L. M.; Zabarnick, S. *Energy and Fuels* **2009**, *23* (11), 5474–5482.
- (24) Rosenfeld, J. M. *Anal. Chim. Acta* **2002**, *465* (1-2), 93–100.
- (25) Li, N.; Ma, X.; Zha, Q.; Song, C. *Energy and Fuels* **2010**, *24* (10), 5539–5547.
- (26) Adams, R. K.; Zabarnick, S.; West, Z. J.; Striebich, R. C.; Johnson, D. W. *Energy and Fuels* **2013**, *27* (5), 2390–2398.
- (27) Anastopoulos, G.; Lois, E.; Karonis, D.; Kalligeros, S.; Zannikos, F. *Energy* **2005**, *30* (2-4 SPEC. ISS.), 415–426.
- (28) Balster, L. M.; Zabarnick, S.; Striebich, R. C.; Shafer, L. M.; West, Z. J. *Energy and Fuels* **2006**, *20* (6), 2564–2571.
- (29) Barrow, M. P.; Peru, K. M.; Headley, J. V. *Anal. Chem.* **2014**, *86* (16), 8281–8288.
- (30) Schiewek, R.; Lorenz, M.; Giese, R.; Brockmann, K.; Benter, T.; Gäb, S.; Schmitz, O. J. *Anal. Bioanal. Chem.* **2008**, *392* (1-2), 87–96.
- (31) Hsu, C. S.; Qian, K.; Chen, Y. C. *Anal. Chim. Acta* **1992**, *264* (1), 79–89.
- (32) Kendrick, E. *Anal. Chem.* **1963**, *35* (13), 2146–2154.
- (33) Hughey, C. A.; Hendrickson, C. L.; Rodgers, R. P.; Marshall, A. G.; Qian, K. *Anal. Chem.* **2001**, *73* (19), 4676–4681.
- (34) Tautenhahn, R.; Bottcher, C.; Neumann, S. *BMC Bioinformatics* **2008**, *9*.
- (35) Kováts, E. *Helv. Chim. Acta* **1958**, *41* (7), 1915–1932.
- (36) Lee, M. L. *Anal. Chem.* **1979**, *51* (6), 768–774.
- (37) Tong, H.; Bell, D.; Tabei, K.; Siegel, M. M. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (11), 1174–1187.
- (38) Keller, B. O.; Sui, J.; Young, A. B.; Whittal, R. M. *Anal. Chim. Acta* **2008**, *627* (1), 71–81.
- (39) Waldo, G. S.; Mullins, O. C.; Penner-Hahn, J. E.; Cramer, S. P. *Fuel* **1992**, *71* (1), 53–57.
- (40) Gong, Y.; Li, Z.; An, J. *J. Dispers. Sci. Technol.* **2005**, *26* (4), 503–507.
- (41) Liu, Z.; Patterson, D. G.; Lee, M. L. *Anal. Chem.* **1995**, *67* (21), 3840–3845.

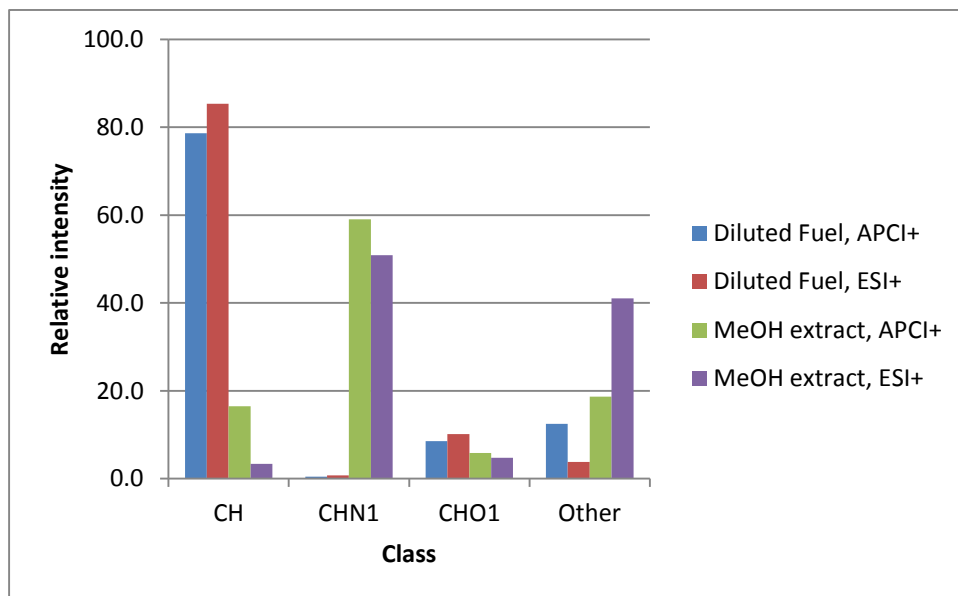
**Supporting material.** Mass spectral peaks that were used for internal calibration of FT-ICRMS data. Specific parameter settings used in FT-ICRMS experiments that may influence ion transmission. Additional results comparing the distribution of classes obtained when diluted fuels and methanol extracts were analyzed with ESI and APCI. This material is available free of charge via the Internet at <http://pubs.acs.org>.

**Table S1.** Mass spectral peaks that were used for the internal calibration of FT-ICRMS data.

ESI-positive		ESI-negative	
<i>m/z</i>	sum formula	<i>m/z</i>	sum formula
134.097	C <sub>9</sub> H <sub>12</sub> N	183.0116	C <sub>8</sub> H <sub>7</sub> O <sub>3</sub> S
160.1126	C <sub>11</sub> H <sub>14</sub> N	205.1229	C <sub>13</sub> H <sub>17</sub> O <sub>2</sub>
176.1439	C <sub>12</sub> H <sub>18</sub> N	221.1542	C <sub>14</sub> H <sub>21</sub> O <sub>2</sub>
202.1596	C <sub>14</sub> H <sub>20</sub> N	261.2218	C <sub>18</sub> H <sub>29</sub> O
256.2065	C <sub>18</sub> H <sub>26</sub> N	281.2481	C <sub>18</sub> H <sub>33</sub> O <sub>2</sub>
312.2691	C <sub>22</sub> H <sub>34</sub> N	301.2168	C <sub>20</sub> H <sub>29</sub> O <sub>2</sub>
438.41	C <sub>31</sub> H <sub>52</sub> N	317.2844	C <sub>22</sub> H <sub>37</sub> O
471.4178	C <sub>30</sub> H <sub>56</sub> O <sub>2</sub> Na	329.2844	C <sub>23</sub> H <sub>37</sub> O
		345.3157	C <sub>24</sub> H <sub>41</sub> O
		367.3576	C <sub>24</sub> H <sub>47</sub> O <sub>2</sub>

**Table S2.** Ion transmission settings for FT-ICRMS data acquisition (n.a. – not available, n.p. not present in the instrument).

Assembly	Part	ESI (solariX)		APCI (apeX)
		(+)	(-)	(+)
Source Optics	Funnel RF Amplitude	40 Vpp	40 Vpp	150 Vpp
Octopole/Hexapole	Frequency	5 MHz	5 MHz	5 MHz
	RF Amplitude	300 Vpp	300 Vpp	400 Vpp
Collision cell	Frequency	2 MHz	2 MHz	n.a.
	RF Amplitude	1000 Vpp	1000 Vpp	800 Vpp
Transfer optics	Time of flight	0.6 ms	0.6 ms	0.7 ms
	Frequency	6 MHz	6 MHz	n.p.
	RF Amplitude	400 Vpp	400 Vpp	n.p.



**Figure S1.** Relative intensities of ion classes observed for diluted fuels and methanol extracts when analyzed with ESCI-TOFMS. Similar parameters as discussed in the Experimental section (ESI-TOFMS) were used. The results show that many species forming part of the CH-class were not extracted efficiently with methanol. For the methanol extract, the CH-class was more prominent in the APCI spectrum, therefore showing the selectivity of the different ionization techniques.