Mapping and quantifying isomer sets of hydrocarbons (≥C$_{12}$) in diesel exhaust, lubricating oil and diesel fuel samples using GC×GC-ToFMS

Mohammed S. Alam[1], Soheil Zeraati-Rezaei[2], Zhirong Liang[3], Christopher Stark[1], Hongming Xu[2], A. Rob MacKenzie[1¥] and Roy M. Harrison[1*†]

[1] **Division of Environmental Health and Risk Management, School of Geography, Earth and Environmental Sciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom**

[2] **Department of Mechanical Engineering, University of Birmingham, Edgbaston Birmingham, B15 2TT, United Kingdom**

[3] **School of Energy and Power Engineering, Beihang University Beijing, 100191 China**

---

[*] To whom correspondence should be addressed.
Tele: +44 121 414 3494; Fax: +44 121 414 3709; Email: r.m.harrison@bham.ac.uk

[†] Also at: Department of Environmental Sciences / Center of Excellence in Environmental Studies, King Abdulaziz University, PO Box 80203, Jeddah, 21589, Saudi Arabia

[¥] Also at: Birmingham Institute of Forest Research, University of Birmingham, B15 2TT

ABSTRACT

Airborne particles and vapours, like many other environmental samples including water, soils, and sediments, contain complex mixtures of hydrocarbons, often deriving from crude oil either before or after fractionation into fuels, lubricants and feedstocks. Comprehensive 2D Gas Chromatography- Time-of-Flight Mass Spectrometry (GC×GC-ToFMS), offers a very powerful technique separating and identifying many compounds in complicated hydrocarbon mixtures. However, quantification and identification of individual constituents at high ionization energies would require hundreds of expensive (when available) standards for calibration. Although the precise chemical structure of hydrocarbons does matter for their environmental impact and fate, strong similarities can be expected for compounds having very similar chemical structure and carbon number. There is, therefore, a clear benefit in an analytical technique which is specific enough to separate different classes of compounds, and to distinguish homologous series, whilst avoiding the need to handle each isomer individually. Varying EI (electron impact) ionization mass spectrometry significantly enhances the identification of individual isomers and homologous compound groups, which we refer to as 'isomer sets'. Advances are reported in mapping and quantifying isomer sets of hydrocarbons ($\geq C_{12}$) in diesel fuel, lubricating oil and diesel exhaust emissions. Using this analysis we report mass closures of *ca.* 90% and 75% for diesel fuel and lubricating oil, and identify 85% and 75% of the total ion current for gas and particulate phase diesel exhaust emissions.

1.   INTRODUCTION

Crude oil contains a highly complex mixture of chemical constituents, mainly hydrocarbons (C_4-C_{55}) (Riazi, 2005). There are many reports of crude oil entering the environment through spillage, or deliberate release (Gertler et al., 2010). Most crude oil is treated and fractionated in order to produce fuels and lubricants for use in transport and combustion applications, and as feedstocks for the chemical industry (Riazi, 2005).   All of these uses have a potential to contaminate the environment. Understanding fates, pathways, and effects of contamination requires chemical analysis and detailed interpretation of resulting data. Much of the chemical complexity of oil derives from the large numbers of straight and branched chain and cyclic hydrocarbon isomers for a given carbon number (Goldstein et al., 2007).   Hence, analytical methods are required that can discriminate structurally similar sets of isomers in complex media.

Application of conventional gas chromatographic methods to oils and oil-derived samples was for many years severely limited by the poor separation capability of one-dimensional chromatography, due to the near-continuous range of physicochemical properties of hydrocarbons.   Thus, typically 90% of the hydrocarbon content of the sample is present in the unresolved complex mixture (UCM), creating a large hump in the chromatogram (Fraser et al., 1998; Schauer et al., 1999).   The advent of two-dimensional gas chromatography, which provides enhanced separating capability due to the orthogonal separation by two capillary columns of different stationary phases, has transformed the problem by resolving the UCM into many thousands of individual compound peaks. The two columns are connected in series by a modulator which is employed to provide focusing of the primary column eluent (Liu et al., 1991; Phillips et al., 1995).   Large amounts of data are produced

72 due to the large number of compounds separated. The information which is most useful

73 scientifically in order to compare the main compositional attributes of samples is often more

74 detailed than that provided by bulk Total Hydrocarbon measurements or fractionation into a few

75 volatility classes, but less detailed than specific identification of many compounds.   Use of a flame

76 ionisation detector has the advantage of allowing generic quantification of any part of the

77 chromatogram in terms of the carbon mass contained within it, but identification of specific

78 chemical constituents with this detector can only be achieved on the basis of retention times which,

79 in a very complex two-dimensional chromatogram and set-up-dependent chromatogram, are

80 laborious to assign objectively. Mass spectrometric detection, especially when employing both low

81 and high ionisation energies adds a third analytical dimension with the ability to overcome the

82 problem of compound identification (Alam et al., 2016a), but has not generally been applied to

83 generic quantification of compound groups within complex samples.   In this study, we show that

84 time-of-flight mass spectrometric detection can be used not only to identify and quantify individual

85 chemical constituents within the chromatogram, but can also be used to quantify generic groups of

86 compounds.

87

88 Motor vehicles are a major source of organic carbon in the atmosphere, and the majority of the fine

89 particulate matter (PM) emitted is carbonaceous, directly emitted as primary organic aerosol (POA)

90 or formed as secondary organic aerosol (SOA) (Jimenez et al., 2009). A substantial fraction of the

91 POA in vehicle emissions has been shown to be semi-volatile under atmospheric conditions

92 (Robinson et al., 2007; May et al., 2013), and is mainly comprised of aliphatic species in the carbon

93 number range between C_{12}–C_{35}, with effective saturation concentrations (C*) between 0.1 and 10^3

94 µg m^{-3} (Robinson et al., 2007; Weitkemp et al., 2007).   The semi volatile organic compound

(SVOC) composition of lubricating oil has been reported to be dominated by branched, cyclic and straight alkanes (≥80%), with the largest contribution from cycloalkanes (≥27%) (Sakurai et al., 2003; Worton et al., 2014).

Previous research has used a limited range of tracer compounds, or homologous series, for the quantification of emissions, considering representative species that can be distinguished from the bulk of the mass, typically involving analysis of the n-alkanes, polycyclic aromatic hydrocarbons (PAH), hopanes, and steranes (Schauer et al., 1999; 2002) each of which represent only a small fraction of the total mass or number of compounds emitted and might lead to underestimation of the importance of lubricating oil as a source of SOA (Brandenberger et al., 2005; Fujita et al., 2007). Rogge et al. (1993) investigated the sources of fine organic aerosol from non-catalyst and catalyst equipped vehicles using one-dimensional GC/MS, but could only resolve 10-15% of the organics, including n-alkanes and PAH. Although some studies have utilized soft ionization to analyse diesel fuel at a molecular level, (Briker et al., 2001); Eschner et al., 2010; Amirav et al., 2008) very few studies have analysed lubricating oil at a molecular level that includes the analysis of SVOCs (Worton et al., 2015; Reddy et al., 2012). In order to address the problems of coelution of constituents of the UCM, Worton et al. (2014) and Isaacman et al. (2012) utilized gas chromatography coupled with vacuum ultraviolet ionization mass spectrometry (GC/VUV-MS) to study the constitutional isomers present in lubricating oil and diesel fuel, respectively, and in a standard crude oil from the Gulf of Mexico (Worton et al., 2015). More recently, Goodman-Rendall et al. (2016) used GC/MS with cold electron impact (EI) ionization, resolving detailed molecular components of diesel fuel. Their results showed that the most important factors in determining SOA yields were carbon number, the presence (or absence) of a ring moiety and the degree of

118   substitution; and precise information of branching and degrees of unsaturation was of secondary

119   importance. Dunmore et al. (2015) have shown that diesel-related hydrocarbons are responsible for

120   60 % of the winter primary hydrocarbon hydroxyl radical reactivity and possibly up to 50 % of the

121   ozone production potential in London. Detailed chemical characterization of diesel emissions would

122   therefore not only resolve factors in determining the contribution to SOA yields, but also to shed

123   light on specific precursors with large photochemical ozone creation potentials and OH reactivity,

124   as well as the identification of compounds that are harmful for human health and the environment.

125

126

127   As well as identifying individual compounds, using GC×GC allows compounds of similar chemical

128   structure to be classified into distinct groups in ordered chromatograms based on their volatility and

129   their polarity, providing information that aids identification and assessment of environmental fate.

130   Dunmore et al. (2015) recently grouped low molecular weight ($\leq C_{12}$) hydrocarbons in atmospheric

131   samples by carbon number and functionality using GC×GC. They reported the grouping of C_6–C_{13}

132   aliphatics and C_2–C_4 substituted monoaromatics, combining the area of all the peaks contained

133   within their selected areas.

134

135   In our study, two dimensional Gas Chromatography Time-of-Flight-Mass Spectrometry

136   (GC×GC-ToF-MS) (Adahchour et al., 2008; Alam et al., 2013; Alam et al., 2016b) was combined

137   with an innovative quantification methodology based on total ion current (TIC) signal response to

138   provide identification and quantification for the compound classes within typical diesel fuel, engine

139   lubricant and engine emissions (gas and particulate phases), providing   a near complete mass

140   closure for diesel fuel and engine lubricant and analyses of diesel engine exhaust composition.

141

2.   EXPERIMENTAL

2.1   Sampling

Gas and particulate phase diesel exhaust emissions were collected from a light-duty diesel engine.

This 2.2 L, 4-cylinder, in-line compression ignition engine was equipped with a common rail direct

injection system and a variable-nozzle-turbine (VNT) turbocharger. Samples were collected with no

diesel oxidation catalyst (DOC) and diesel particulate filter (DPF). The diesel engine emissions

were diluted (1:50) with cleaned compressed air using an in-house exhaust dilution system. Samples

were collected at steady state engine operating conditions at a low engine load of 3.0 bar mean

effective pressure (BMEP) and a speed of 1800 revolutions per minute (RPM). The temperature at

the sampling point was 25 ± 5 °C and samples were collected for 30 min at a flow rate of 1.8 L

min^{-1}. Adsorption tubes were used to collect gas phase constituents directly from the diluted diesel

engine exhaust, downstream of a polypropylene backed PTFE 47 mm filter (Whatman, Maidstone,

UK), which was used to collect constituents in the particulate phase. Further details of the engine

exhaust sampling system are given elsewhere (Alam et al. 2016c).

156

Diesel fuel, engine lubricating oil, and gas-and-particulate diesel exhaust emission samples were

analysed using GC×GC-ToF-MS. Briefly, 1 μL of diesel fuel (EN 590-ultra low sulfur diesel, Shell,

UK) was diluted (1:1000) in dichloromethane (DCM) and injected onto a stainless steel thermal

adsorption tube, packed with 1 cm quartz wool, 300 mg Carbograph 2TD 40/60 (Markes

International, Llantrisant, UK), for analysis on the thermal desorber (TD) coupled to the

GC×GC-ToF-MS. The EN 590 ultra-low sulphur diesel fuel is representative of the standardized

ultra-low sulphur content fuel (<10 mg kg^{-1} or ppm) that is widely used in the UK and Europe

(European Parliament and the Council of the European Union, 2009). 1 µL of engine lubricant

(fully synthetic, 5W30, Castrol, UK) was diluted (1:1000) in DCM and directly injected into the gas

chromatographic column, as the high molecular weight constituents found in the lubricating oils

would not efficiently desorb into the GC column from the adsorption tubes.

2.2   Instrumentation

Adsorption tubes were desorbed using TD (Unity 2, Markes International, Llantrisant, UK) and

samples were subsequently analysed using a gas chromatograph (GC, 7890B, Agilent Technologies,

Wilmington, USA) equipped with a Zoex ZX2 cryogenic modulator (Houston, USA). The primary

column (first separation dimension) was equipped with a SGE DBX5, non-polar capillary column

(30 m, 0.25 mm ID, 0.25 µm – 5% phenyl polysilphenylene-siloxane). The secondary, more polar

column (second separation dimension) was equipped with a SGE DBX50 (4.0 m, 0.1 mm ID, 0.1

µm – 50% phenyl polysilphenylene-siloxane), situated in a secondary internal oven. The GC×GC

was interfaced with a BenchTOF-Select, time-of-flight mass spectrometer (ToF-MS, Markes

International, Llantrisant, UK), with a scan speed of 50 Hz and a mass resolution of >1200 full

width at half maximum (fwhm) at 70 eV and >800 fwhm at 14eV over 100 – 1000 m/z. The

mass/charge range was 30 to 525 m/z, and quantification was conducted on nominal unit mass

resolution. Electron impact ionisation energies on this ToFMS can be tuned between 10 eV and 70

eV, the former retaining the molecular ion, while the latter causes extensive fragmentation, but

allows comparison with standard library spectra (Alam et al. 2016a). Data were processed by using

GC Image v2.5 (Zoex Corporation, Houston, USA).

2.3   Standards & Chromatography Methodology

Nine deuterated internal standards namely, dodecane-d_{26}, pentadecane-d_{32}, eicosane-d_{42}, pentacosane-d_{52}, triacontane-d_{62}, biphenyl-d_{10}, n-butylbenzene-d_{14}, n-nonylbenzene-2,3,4,5,6-d_5 (Chiron AS, Norway) and p-terphenyl-d_{14} (Sigma Aldrich, UK) were utilised for quantification. Natural standards included 24 n-alkanes ($C_{11} - C_{34}$), phytane and pristane (Sigma Aldrich, UK), 16 n-alkylcyclohexanes ($C_{11} - C_{25}$ and C_{26}), 5 n-alkylbenzenes (C_4, C_6, C_8, C_{10} and C_{12}), cis- and trans-decalin, tetralin, 4 alkyltetralins (methyl-, di-, tri- and tetra-), 4 n-alkyl naphthalenes (C_1, C_2, C_4 and C_6) (Chrion AS, Norway) and 16 USEPA polycyclic aromatic hydrocarbons (Thames Restek UK Ltd). These standards were chosen in order to cover as much of the overall chromatogram as possible. The chromatography methodology (*i.e.* GC oven temperatures, temperature ramp rates *etc.*) of the analysis of the adsorption tubes, lubricating oil and gas/particulate phase samples is discussed in Section S1 in the Supplementary Information.

2.4   Grouping of Chromatographically Resolved Compounds

Structurally similar compounds possess similar physicochemical properties. This facilitates identification when separating a mixture according to these physical and chemical properties. Diesel fuels, diesel emissions and lubricating oils have been shown to consist of a limited number of compound classes, but an enormous number of individual components within a class.

In this study we use GC×GC coupled to variable ionisation ToF-MS to map and quantify isomer sets previously unresolved in the UCM. Conventional electron ionisation at 70 eV imparts a large amount of excess energy causing extensive fragmentation, with a tendency to generate similar mass

210 spectra. Thus for example the isomeric alkanes all exhibit the same m/z 43, 57, 71, 85, 99 patterns,

211 thus obscuring the match with the NIST library and making identification from the mass spectrum

212 very difficult. To address this issue, a lower ionization energy (10-14eV) was also employed in our

213 study so that the organic compounds are ionized with minimal excess internal energy and thus less

214 fragmentation, hence retaining the distinct identity of the molecule with a much larger fraction of

215 the molecular ion (Alam et al., 2016a).   Running samples on the GC×GC with both low and high

216 ionization energy mass spectrometry results in a wealth of data for identification of compounds;

217 where 14 and 70eV mass spectra can be compared for a given species owing to the identical

218 retention times of the repeat runs. At low ionization energy the molecular ion is enhanced, at the

219 same time as retaining some fragmentation, while at high ionization energy the mass fragmentation

220 patterns of a species can be compared directly to mass spectral libraries. This allows easier

221 identification of unknown compounds. Low EI spectra therefore give qualitative information, while

222 high EI mass spectra allow for quantitative analyses to be performed.

223

224 Our recent work exploited soft ionisation (14 eV) to identify a large number of isomers,

225 demonstrating the ability to separate and identify individual alkanes (normal, branched and cyclic)

226 with specific carbon numbers, based on their volatility and polarity (Alam et al., 2016a).   In this

227 study we expand our previous qualitative analysis and separate the alkane series (as well as other

228 homologous series) into isomer sets containing the same carbon number. Individual alkanes that

229 were identified as having different molecular ions (i.e. different carbon number) to the *n*-alkane

230 within the area of the chromatogram were included in their appropriate adjacent (usually *n*±1) area;

231 for example, some dimethyl isomers can be shifted by ~100 delta-Kovats (~1 carbon number),

232 whereas trimethyl- and tetramethyl isomers have been reported to be shifted by ~150 and ~200

delta-Kovats. This has been completed for all the homologous series reported in this study. The grouping of the alkanes according to their respective carbon numbers is shown in Figure 1, where the least polar compounds (fast eluting peaks in the second dimension) are the alkanes, increasing in carbon number as the retention time in the first dimension increases.

This methodology was expanded to more polar homologous series including monocyclic alkanes, bicyclic alkanes, tricyclic alkanes, tetralins/indanes, monocyclic aromatics, bicyclic aromatics and alkyl-biphenyls. Like the alkanes, a significant problem in creating the boundaries of groups is the overlapping of one carbon number group into another. Identifying each individual compound in this case (as with the alkanes above) would be resource and time intensive and so carefully constructed 'Computer Language for Identification of Chemicals' (CLIC) qualifiers were created and utilised in order to match peaks and their mass fragmentation patterns. A CLIC qualifier is an expression in a computer language that allows users of chromatographic software to build rules for selecting and filtering peaks using retention times and mass fragmentation patterns (Reichenbach et al., 2005). This was exploited to identify specific compounds belonging to a compound class and a polygon selection tool within the GC Image software was drawn around this section of the chromatogram (coloured polygons shown in Figure 1). Any overlap in the graphics was accounted for by forcing peaks to belong to one compound class over another via strict mass fragment and molecular ion selection tools. Examples of CLIC expressions utilised for identifying compound classes are included in the Supplementary Information (Section S2). An example of a selected ion chromatogram with a specific CLIC expression is shown in Figure 2, for C_6-substituted monocyclic aromatics, with their corresponding 70eV and 12eV mass spectra. The characteristic 70eV mass fragments at m/z 92, 105, 119, 133 signify cleavage of the C-C bond next to the benzene ring. The

256 12eV mass spectra, however, produce poor characteristic fragment ions, but a prominent molecular

257 ion (162) and m/z 92 signifying the overall mass of the molecule and the benzene ring (Ph-CH$_2^+$),

258 respectively. In effect, the polygons mark out sets of isomeric compounds having the same

259 empirical formula and shared structural elements; the sets appear to intersect each other in the

260 two-dimensional chromatogram space, but compounds in the intersecting regions are assigned

261 uniquely to a class using a third, mass spectrometric, data dimension (i.e. mass fragmentation

262 patterns). The resulting isomer sets are more chemically and environmentally meaningful than the

263 raw polarity/volatility assignment from the chromatography. This approach was completed

264 independently for diesel fuel, lubricating oil, and gas/particulate phase exhaust emissions to ensure

265 the applicability of polygon boundaries and reproducibility of retention times and mass fragments.

266 Results indicated that isomers within the constructed polygon boundaries possessed identical

267 retention times and interpretable mass spectra for all differing samples. Retention times were

268 reproducible in both chromatographic dimensions in separate runs. The isomer sets (polygons) can

269 be linked together in a large template that can also be linked to the internal standards. This allows

270 easy alignment of the template (all isomer sets) in the event of slight shifts in retention times, due to

271 column changes or instrumental maintenance.

272

273 **2.5   Quantification of compounds with no authentic standards**

274 The authentic standard mixture contains 74 standard compounds including 9 internal standards (see

275 Section 2.3). These standards were chosen in order to cover as much of the overall chromatogram as

276 possible and are used for obtaining a calibration for quantifying groups of isomers with the same

277 molecular ion and functionality. For example; the response for C$_{11}$ (undecane, m/z 156) was used to

278 quantify all C$_{11}$ alkane isomers which were positively identified in the analysed samples (and have

279 the same molecular mass and retention times in all sample runs). Using retention times for C_{11}

280 alkane isomers as well as mass spectra, only those C_{11} isomer peaks were selected that were

281 observed using the SIC for m/z 156 (or the respective CLIC expressions). Isomer sets were

282 comprehensively identified in the analysed samples using mass fragmentation at 70eV and

283 molecular masses at lower ionisation energies (10-15eV). The total ion current within each polygon

284 was integrated and the isomer set abundance was estimated using the response ratio of the closest

285 structurally-related deuterated standard to the corresponding compound class natural standard with

286 the same carbon number (usually within the polygon). This methodology has an uncertainty of

287 approximately 24% and is discussed in detail in the Supplementary Information (Section S3).

288

289 **3. RESULTS AND DISCUSSION**

290 **3.1   Analysis of Diesel Fuel**

291 The chromatography of the diesel fuel analysed by TD-GC×GC-ToF-MS is shown in Figure 1.

292 Compounds identified within the diesel fuel included: *n*-alkanes, branched alkanes (mono-, di-, tri-,

293 tetra- and penta-methyl), *n*-alkyl cycloalkanes, branched monocyclic alkanes, C_1-C_{12} substituted

294 bicyclic alkanes, C_1-C_4 substituted tetralins and indanes, C_3-C_{12} substituted monocyclic aromatics,

295 C_1-C_3 substituted biphenyls/acenaphthenes, C_1-C_4 substituted bicyclic aromatics, C_1-C_2 substituted

296 fluorenes (FLU), C_1-C_2 substituted phenanthrene/anthracenes (PHE/ANT) and unsubstituted PAH.

297 Representative mass spectra at 12eV and 70eV ionization are presented in the Supplementary

298 Information (Section S4). These compounds accounted for 93% of the total response (excluding the

299 siloxanes which derive from contaminants *i.e.* column bleed) and was equivalent to 90 % of the

300 mass injected. Therefore, out of the 8026 (± 24) ng that was injected into the GC×GC (mass

301 calculated by weighing and diluting a known volume of diesel fuel, while the uncertainty represents

the combined uncertainties of the processes involved in this estimation) a mass of approximately

7200 (± 1728) ng was accounted for. We suspect that a significant amount of the mass that was

unaccounted may be $<C_{10}$ and/or any unresolved peaks that we were unable to measure and/or

identify using our technique. The percentage contribution of each compound class identified to the

total mass accounted for is shown in Table 1.

Our results indicate that the majority of the diesel fuel consists of aliphatic compounds, with a low

aromatic content (~10%). Very few published studies exist elucidating the contribution of different

constituents in diesel fuel (Isaacman et al., 2012; Welthagen et al., 2007; Gentner et al., 2012).

Most studies focus on the characterisation of specific compounds within diesel fuels such as

nitrogen containing species (Wang et al. 2004), cyclic compounds (Edam et al., 2005), or to identify

strengths and weaknesses in analytical techniques (Frysinger et al., 1999). Recently, VUV

ionization at $10 - 10.5$ eV has been exploited to elucidate some of the structures within diesel fuel,

by separating components using GC (Isaacman et al., 2012). The authors report their observed

mass of diesel fuel as 73% aliphatic and 27% aromatic, broadly consistent with the results of this

study. Up to 11% of the observed mass fraction of diesel fuel was attributed to bicyclic alkanes, a

factor of 2 larger than observed in this study. Their observed mass fractions of cycloalkanes and

benzene, however, are in excellent agreement. The contribution of branched alkanes (*i*-alkanes) and

linear *n*-alkanes to the total mass of the alkanes was 39.1 and 23.1%, respectively. A significant

proportion of the total mass observed was attributed to alkanes (62%), a factor of 1.5 larger than

reported by Isaacman et al. (2012). However, the differences observed between diesel fuel

analysed in this study and that reported by Isaacman et al. (2012) are attributable to different fuel

formulations and/or fuel source, as opposed to analytical methods. Although not shown here, a significant number of alkane isomers were identified for each carbon number using soft ionisation mass spectrometry, accounting for a total of ~200 alkanes across the $C_{11} - C_{30}$ range. The ratio of i-alkanes to n-alkanes sharply decreases after C_{25}, indicating a reduced amount of mass represented by branched isomers present in diesel fuel for $>C_{25}$ alkanes, which could be related to the formulation process, or reflect the composition of the feedstock.

3.2   Analysis of diesel engine emissions (gas phase)

A GC×GC contour plot of the gas phase diesel exhaust emissions is shown in the Supplementary Information (Figure S5). The observed chromatogram for the gas phase emissions looked extremely similar to that of the diesel fuel chromatogram (Figure 1), suggesting that the majority of compounds found in the gas phase emissions are of diesel fuel origin. All of the compounds found in the diesel fuel were observed in the gas phase emissions, albeit with a reduced number of i-alkanes $>C_{20}$, which may signify efficient combustion of these high molecular weight compounds, or partitioning into the particulate phase. The measured constituents of the gas phase diesel exhaust emissions are shown in Table 1. Approximately 15 % of the total ion current (response, excluding siloxanes) was unaccounted for. Table 1 illustrates the percentage mass of each compound class identified, in the 85% of the response that was accounted for. As the total mass of the gas phase sample is unknown, a mass for the remaining 15% of the total response cannot be estimated, as the individual components that are unidentified will have different responses per unit mass. For example, 23.5% of the mass identified was attributed to $C_4 - C_{12}$ alkyl substituted monocyclic aromatics and accounted for 9.7% of the total ion current response; and 10.0% of the mass identified was bicyclic alkanes, representing 9.2% of the response.

349 Although the diesel fuel constituents present in the gas phase exhaust emissions broadly were

350 compositionally consistent with the fuel, there were significant differences observed in their relative

351 amounts. Of the total mass identified in the gas phase emissions, n-alkanes and i-alkanes

352 represented 9.8% and 30.1%, respectively. These are factors of 2.4 and 1.3 lower than that for diesel

353 fuel, respectively; which may be due to preferred combustion of these compounds (Burcat et al.,

354 2012). Enhancements of monocyclic aromatics, monocyclic alkanes, bicyclic alkanes and bicyclic

355 aromatics were observed in the emissions, possibly due to them being intermediate species formed

356 during the combustion of larger molecules (Gentner et al., 2013), and unlikely to be a contribution

357 from lubricating oil as very little mass was attributed to compounds with $<C_{18}$ (see Section 3.3). A

358 very limited number of oxygenates were also identified (*e.g.* ketones (m/z 58, 72), carboxylic acids

359 (m/z 60)), most probably combustion products of diesel fuel, but representing a very small fraction

360 of the total measured gas phase emissions (<1%). However, carboxylic acids are difficult to detect

361 using GC/MS without prior derivatization and may therefore be underestimated. Gentner et al.

362 (2013) suggest that compounds such as alkenes, aromatics and oxygenates comprise ~30% of the

363 total measured gas phase emissions, in agreement with this study; however, they suggest that these

364 products are unlikely to contribute to primary organic aerosol (POA). We observe these aromatic

365 compounds in the particulate phase also, indicating a contribution.

366

367 **3.3   Analysis of Lubricating Oil**

368 The isomer sets (polygons), identified in the lubricating oil chromatogram, are shown in Figure 3A.

369 Molecular ions present in the mass spectra enabled the grouping of isomers by carbon number,

370 while the presence of the characteristic mass fragments, presented in Table 1, were used to confirm

371 the identity of the type of hydrocarbon. The representative mass spectra for compounds presented in

372 Table 1, for the lubricating oil is shown in the Supplementary Information, Section S6. Polygons

373 were drawn around groups of compounds that possessed the same molecular ion for a given

374 compound class, see Figure 3A and 3B. The lubricating oil was analysed using two independent

375 temperature ramps of the GC×GC (methodologies outlined above); one to achieve the best possible

376 comprehensive separation of compounds in the oil (Figure 3A) and the other using methodologies

377 developed for analysis of the particulate phase components of engine exhaust (Figure 3B), to

378 ascertain where the compounds identified in the oil are present in the particulate phase emissions

379 filter. Figure 3B also illustrates the positioning of the SVOC measured in the gas phase, that are

380 observed in the particulate phase filter as well as the positioning for the PAH. The grouping

381 template that is illustrated in Figure 1 covers the SVOC range indicated in Figure 3B.

382

383

384 Using the signature mass fragment ions (Table 1) together with the calculated molecular mass,

385 specific compounds with the same carbon number were isolated, see Figure S7-A and S7-B. For

386 example, selecting the ion fragment m/z 92 and 119 for monocyclic aromatics gives rise to the

387 selected ion chromatogram illustrated in Figure S7-A. This can be achieved using 70eV mass

388 spectrometry identifying a homologous series across a large carbon number range. However,

389 selecting the molecular mass for a specific carbon number allows the identification of all isomer

390 sets in a region of the chromatogram with that specific molecular mass, as shown in Figure S7-B for

391 C_{22} monocyclic aromatics (m/z 302). A mass of 8511 (±255) ng of lubricating oil was injected into

392 the GC×GC, of which 6356 (±1525) ng was quantified. This methodology was used to identify and

393 quantify the following homologous series: C_{16}–C_{33} straight and branched chain alkanes, C_{16}–C_{33}

394 monocyclic alkanes, C_{17}–C_{33} bicyclic alkanes, C_{17}–C_{33} tricyclic alkanes and C_{16}–C_{33} monocyclic

395 aromatics. These compound groups represented approximately 91% of the total ion current

396 (excluding siloxanes) and 75% of the mass fraction. Adamantanes, diamantanes, pentacyclic and

397 hexacyclic alkanes, steroids, steranes, and hopanes represented 5% of the total ion current, while the

398 remaining 4% remained unidentified. These compounds were not quantifiable using this

399 methodology, as there were no standards available that corresponded to these sections of the

400 chromatography and could not be estimated as they are not present in a homologous series.

401 However, from previous literature, these compound classes are thought to represent a small fraction

402 of the mass (Worton et al., 2015). Furthermore, we have not taken into account any

403 non-organic/hydrocarbon species but according to these data, the fraction of any non-HCO material

404 is likely to be small.

405

406

407 Worton et al. (2015) exploited VUV photoionization mass spectrometry to characterize

408 comprehensively hydrocarbons in a standard reference crude oil sample. They reported a total mass

409 closure of $68 \pm 22\%$, comprised of linear and branched alkanes (19%), 1-6 ring cycloalkanes (37%),

410 monoaromatics (6.8%) and PAH (4.7%). The mass fractions observed for linear and branched

411 alkanes in this study were 11% and 12%, respectively, which is in excellent agreement. There is

412 also excellent agreement with the mass attributed to bicyclic (2-ring) and tricyclic (3-ring) alkanes

413 however, for monocyclic alkanes the results presented here are a factor of 2 larger than Worton et

414 al.(2015) and 2.5 larger than Reddy et al. (2012). Both previous studies analysed similar crude oil

415 samples associated with the Deepwater Horizon disaster (McNutt et al., 2012), and would be

416 expected to differ appreciably from a lubricating oil. Furthermore, no PAH were observed in the

lubricating oil in this study, in agreement with Zielinska et al. (2004) but in contrast to Worton et al. (2015). We attribute this difference to the varying crude oil origins and formulation processes involved in the production of synthetic oil.

Previous work from this group identified a large number of isomeric species in base oil using 14eV EI ionization energy mass spectrometry (Alam et al., 2016a). Although we were able to identify a large number of compounds, there still existed a small amount of fragmentation at 14eV, particularly for alkyl-methyl-, alkyl-dimethyl-, and alkyl-trimethyl-cyclohexanes. In this study the fragmentation was significantly reduced for these compounds at 12eV (i.e. relative intensities of *m/z* 97, 111, 125 reduced by >50%) and completely eradicated (relative intensities of mass fragments reduced by >95%) at 10eV, leaving the *m/z* 82 ion (for monocyclic alkanes) and the molecular ion. This demonstrates the significant differences observed in fragmentation over small changes in lower ionisation EI energies and may also account for slight discrepancies between studies (Worton et al., 2015; Isaacman et al., 2012; Alam et al., 2016b).   Utilising these differences in fragmentation from using low ionization energies (10 – 15eV) may provide more information in regards to the structure of many isomeric compounds.

3.4   Analysis of Diesel Engine Emissions (Particulate Phase)

90% of the total ion current of the particulate phase filter was identified and attributed to a wide range of classes. Of the total mass identified, 47 (±11)% was straight and branched chain alkanes, 20 (±4.8)% monocyclic alkanes, 7.5 (±1.8)% bicyclic alkanes, <3 (±0.7)% tricyclic alkanes, 6 (±1.4)% monocyclic aromatics, 7 (±1.7)% oxygenates, <1 (±0.2)% furanones, 4 (±1.0)% PAH and 2 (±0.5)% fatty acid methyl esters (FAMES). Figure 4 illustrates the percentage mass contribution

440 of homologous series (including isomers) identified as a function of carbon number. Peak

441 concentrations of alkanes (cyclic and straight/branched chain) were observed between $C_{24} - C_{27}$

442 consistent with the lubricating oil, while a small peak in concentration was also observed in the C_{15}

443 $- C_{20}$ range, consistent with the fuel and gas phase emissions. Oxygenated compounds were found

444 to be present in the $C_{11} - C_{22}$ range, suggesting that these compounds are combustion products. The

445 concentration of monocyclic aromatics was steady throughout the carbon number distribution ($C_{15} -$

446 C_{32}), with a small peak at $C_{25} - C_{27}$. The presence of PAH in the particulate phase suggests their

447 formation via diesel fuel combustion or unburnt fuel, owing to their absence in the lubricating oil.

448 There are numerous studies reporting the absence of PAH in unused lubricating oil and presence in

449 used oil, which suggests the absorption of blow-by exhaust containing fuel combustion associated

450 PAH (Fujita et al., 2006). FAMES were identified by their characteristic fragmentation at 70eV EI

451 ionization and with their characteristic ion (m/z 174) at low EI ionization (12 eV).

452

453

454 There have been few studies investigating the contribution of lubricating oil and fuel to the emitted

455 diesel POA, suggesting 20 to 80% influence from lubricating oil (Worton et al., 2014;

456 Brandenberger et al., 2005; Kleeman et al., 2008; Sonntag et al., 2012).   Most recently, it has been

457 suggested that ≥80% of the SVOC composition is dominated by branched cycloalkanes with one or

458 more rings and one or more branched alkyl side chain (Worton et al., 2014). This is significantly

459 larger than that reported in this study (≥30 %), where the majority of the emissions are dominated

460 by straight and branched chain alkanes (47%) over a volatility range that also suggests a significant

461 contribution from the diesel fuel ($C_{11} - C_{20}$, see Figure 4). The diesel fuel and lubricating oil

462 contained respectively 62% and 47.5% straight and branched chain alkanes (summed), suggesting a

larger possible contribution of diesel fuel to the vapour phase engine emissions (which is dominated

by straight and branched chain alkanes). The contribution of unburned lubricating oil, however,

most likely dominates the SVOC emissions in the particulate phase, as shown in Figure 4.

5.   CONCLUSION

The SVOC contents in diesel fuel, 5W30 synthetic lubricating oil, and diesel exhaust emissions

(both in the gas and particulate phases) were characterized using TD-GC×GC-ToFMS. By

exploiting the mass spectrometric fingerprint of eluting compounds in highly structured and ordered

chromatograms, a method has been constructed that quantifies the contributions of 'isomer sets' (i.e.

structural isomers in specific compound classes) to the overall composition of a sample. We found

that the ion current for identified homologous series exhibited very similar responses, illustrating

that quantitative calibrations derived from the *n*-alkane series could be used to estimate the

concentrations of isomeric aliphatic compounds with similar molecular weight. Using this

methodology together with a range of standards, and aggregating compound classes of similar

functionality together (i.e. *n*-alkanes, branched alkanes etc.), we present comprehensive

characterization of diesel fuel, lubricating oil and diesel exhaust emissions.

Furthermore, combining conventional 70eV EI ionization mass spectrometry with lower ionization

energy (10-14eV), allowed the identification of constitutional isomers of the same molecular weight

and compound class, enabling a clear distinction between carbon number and functionality. By

utilising this innovative method, a number of findings were achieved. 1) A mass closure accounting

for *ca.* 90 % and 75 % for the analysis of diesel fuel and lubricating oil, respectively. 2) Acyclic and

monocyclic alkanes were found to be predominant in both the diesel fuel and synthetic lubricating oil (76% and 68%, respectively). 3) Diesel exhaust emissions in the gas phase were extremely similar to the composition of diesel fuel. 4) Diesel exhaust emissions in the particulate phase were similar to the composition of lubricating oil. 5) The presence of combustion products of diesel fuel (e.g. aromatics and oxygenates) in the particulate phase indicates a contribution to POA.

Diesel exhaust hydrocarbons are a significant precursor of secondary organic aerosol (Dunmore et al., 2015; Gentner et al., 2012).   Diesel fuel and lubricant, contributors to diesel exhaust, contain large numbers of isomers.   Separation into isomer sets improves our understanding of the fates of these oil-derived materials in the environment (Lim & Ziemann, 2009; Kroll & Seinfeld, 2008). By utilizing GC×GC-ToFMS with soft ionisation, we enable the identification of the composition of the UCM, characterising the chemical composition by carbon number and compound class, and the possibility of branching structural information. Along with a grouping methodology using CLIC expressions and unique compound fragmentation patterns, we demonstrate the reliable quantitative integration of structural isomers. These methods exploit the improved resolution and isomer separation capabilities of the advanced instrumentation and have potential applications to the observations of petroleum degradation, and SOA formation and evolution. This method is ideal for investigating fossil fuel sources (*e.g.* lubricating oil, motor vehicles, fuel *etc.*) and can be extended to atmospheric measurements where there exist many oxygenates. The analysis of atmospheric particulate phase samples (not using thermal desorption), may benefit from exploring different extraction solvents (e.g. methanol or a hexane/methanol mix) to allow the larger expected proportion of oxygenated species to be efficiently extracted. Although the presence of oxygenated species adds chromatographic complexity, as co-elution can be a limitation, using carefully

constructed CLIC expressions and mass fragmentation patterns, various oxygenates can be identified (*e.g.* 2-ketones (m/z 58), 3-ketones (m/z 72), carboxylic acids (m/z 60) *etc.*), making this technique applicable in any scientific field that routinely characterizes complex hydrocarbon mixtures. Although not explored here, this technique of low ionization mass spectrometry may benefit from reducing the temperature of the ionisation source, as demonstrated by Isaacman et al. (2012); who retained a larger fraction of the molecular parent ion on reducing the ion source temperature of their VUV photoionization technique. This would particularly aid identification for atmospheric samples which would contain more oxygenated compounds.

ACKNOWLEDGEMENTS

The authors would like to thank L. McGregor and N. Bukowski on valuable discussions on aspects of GC×GC-ToFMS and assistance in CLIC qualifiers. Financial support from the European Research Council (ERC-2012-AdG, Proposal No. 320821) for the FASTER project is gratefully acknowledged. The assistance of Y. Al-Qahtani in collecting samples from the engine facility is also acknowledged.

Author Contribution

MSA prepared the manuscript with contributions from ARMK and RMH; RMH, MSA and SZR designed the engine experiments; SZR, MSA and ZL carried out the engine experiments; MSA and CS developed the GC×GC methodology and completed subsequent analyses. HX overlooked the engine facility and RMH oversaw the entire project.

532

Conflict of Interests

534 The authors declare no competing financial interest.

535

REFERENCES

Adahchour, M., Beens, J., & Brinkman, U. A.: Recent developments in the application of comprehensive two-dimensional gas chromatography, Journal of Chromatography A, 1186, 67-108, 2008.

Alam, M. S., West, C. E., Scarlett, A. G., Rowland, S. J., & Harrison, R. M.: Application of 2D-GCMS reveals many industrial chemicals in airborne particulate matter, Atmospheric Environment, 65, 101-111, 2013.

Alam, M. S., Stark, C., & Harrison, R. M.: Using Variable Ionization Energy Time-of-Flight Mass Spectrometry with Comprehensive GC×GC to Identify Isomeric Species, Analytical Chemistry, 88, 4211-4220, 2016a.

Alam, M. S. & Harrison, R. M.: Recent advances in the application of 2-dimensional gas chromatography with soft and hard ionisation time-of-flight mass spectrometry in environmental analysis, Chemical Science, 7, 3968-3977, 2016b.

Alam, M. S., Zeraati-Rezaei, S., Stark, C. P., Liang, Z., Xu, H., & Harrison, R. M.: The characterisation of diesel exhaust particles - composition, size distribution and partitioning, Faraday Discussion, 189, 69-84, 2016c.

Amirav, A., Gordin, A., Poliak, M., & Fialkov, A. B.: Gas chromatography-mass spectrometry with supersonic molecular beams, Journal of Mass Spectrometry, 43, 141-163, 2008.

Brandenberger, S., Mohr, M., Grob, K., & Neukom, H. P. Contribution of unburned lubricating oil and diesel fuel to particulate emission from passenger cars, Atmospheric Environment, 39, 6985-6994, 2005.

Briker, Y., Ring, Z., Iacchelli, A., McLean, N., Rahimi, P. M., & Fairbridge, C.: Diesel fuel analysis by GC-FIMS: Aromatics, n-paraffins, and isoparaffins, Energy and Fuels, 15, 23-37, 2001.

Burcat, A., Dixon-Lewis, G., Frenklach, M., Hanson, R. K, Salimian, S., Troe, J., Warnatz, J., & Zellner, R. Combustion chemistry. Edited by Gardiner, W.J., ed., 2012

Directive 2009/30/EC of the European Parliament and of the Council of 23 April 2009. Official Jounral of the European Union, 2009.

Dunmore, R. E., Hopkins, J. R., Lidster, R. T., Lee, J. D., Evans, M. J., Rickard, A. R., Lewis, A. C., & Hamilton, J. F.: Diesel-related hydrocarbons can dominate gas phase reactive carbon in megacities, Atmospheric Chemistry & Physics, 15, 9983-9996, 2015.

Edam, R., Blomberg, J., Janssen, H.-G., & Schoenmakers, P.: Comprehensive multi-dimensional chromatographic studies on the separation of saturated hydrocarbon ring structures in petrochemical samples, Journal of Chromatography A, 1086, 12-20, 2005.

581 Eschner, M. S., Welthagen, W., Groger, T. M., Gonin, M., Fuhrer, K., & Zimmermann, R.:
582 Comprehensive multidimensional separation methods by hyphenation of single-photon ionization
583 time-of-flight mass spectrometry (SPI-TOF-MS) with GC and GC × GC, Analytical Bioanalytical
584 Chemistry, 398, 1435-1445, 2010.

586 European Parliament and Council of the European Union: Directive 2009/30/EC of the European
587 parliament and of the council of 23 April 2009, Official Journal of the European Union, L140,
588 88-113, 2009

590 Fraser, M. P., Cass, G. R., & Simoneit, B. R.: Gas-phase and particle-phase organic compounds
591 emitted from motor vehicle traffic in a Los Angeles roadway tunnel, Environmental Science &
592 Technology, 32, 2051-2060, 1998.

594 Frysinger, G. S., & Gaines, R. B.: Comprehensive Two-Dimensional Gas Chromatography with
595 Mass Spectrometric Detection (GC×GC/MS) Applied to the Analysis of Petroleum, Journal of High
596 Resolution Chromatography, 22, 251-255, 1999.

598 Fujita, E. M, Campbell, D. E, & Zielinska, B.: Chemical analsis of lubricating oil samples from a
599 study to characterize exhaust emissions from light-duty gasoline vehicles in the Kansas City
600 Metropolitan area, Final Report, 2006.

602 Fujita, E. M., Zielinska, B., Campbell, D. E., Arnott, W. P., Sagebiel, J. C., Mazzoleni, L., Chow, J.
603 C., Gabele, P. A., Crews, W., & Snow, R.: Variations in speciated emissions from spark-ignition and
604 compression-ignition motor vehicles in California's south coast air basin, Journal of Air & Waste
605 Management Association, 57, 705-720, 2007.

607 Gentner, D. R., Isaacman, G., Worton, D. R., Chan, A. W. H., Dallmann, T. R., Davis, L., Liu, S.,
608 Day, D. A., Russell, L. M., Wilson, K. R., Weber, R., Guha, A., Harley, R. A., & Goldstein, A. H.:
609 Elucidating secondary organic aerosol from diesel and gasoline vehicles through detailed
610 characterization of organic carbon emissions. Proceedings of the National Academcy of Science,
611 109, 18318-18323, 2012.

613 Gentner, D. R., Worton, D. R., Isaacman, G., Davis, L. C., Dallmann, T. R., Wood, E. C., Herndon,
614 S. C., Goldstein, A. H., & Harley, R. A.: Chemical composition of gas-phase organic carbon
615 emissions from motor vehicles and implications for ozone production, Environmental Science &
616 Technology, 47, 11837-11848, 2013.

618 Gertler, C, Yakimov, M. M, Malpass, M. C, & Golyshin, P. N.: Shipping-related accidental and
619 deliberate release into the environment, Handbook of Hydrocarbon and Lipid Microbiology,
620 243-256, 2010.

622 Goldstein, A. H., & Galbally, I. E.: Known and unexplored organic constituents in the Earth's
623 atmomsphere, Environmental Science & Technology, 41, 1514-1521, 2007.

625 Goodman-Rendall, K.A., Zhuang, Y. R., Amirav, A., & Chan, A. W. H.: Resolving detailed

molecular structures in complex organic mixtures and modeling their secondary organic aerosol formation, Atmospheric Environment, 128, 276-285, 2016.

Isaacman, G., Wilson, K. R., Chan, A. W. H., Worton, D. R., Kimmel, J. R., Nah, T., Hohaus, T., Gonin, M., Kroll, J. H., Worsnop, D. R., & Goldstein, A. H.: Improved resolution of hydrocarbon structures and constitutional isomers in complex mixtures using gas chromatography-vacuum ultraviolet-mass spectrometry, Analytical Chemistry, 84, 2335-2342, 2012.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Y. M. Zhang, Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., & Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, Science., 326, 1525-1529, 2009.

Kleeman, M. J., Riddle, S. G., Robert, M. A., & Jakober, C. A.: Lubricating oil and fuel contributions to particulate matter emissions from light-duty gasoline and heavy-duty diesel vehicles, Environmental Science & Technology, 42, 235-242, 2008.

Kroll, J. H., & Seinfeld, J. H. Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, Atmospheric Environment, 42, 3593-3624, 2008.

Lim, Y. B., & Ziemann, P. J.: Effects of molecular structure on aerosol yields from OH radical-initiated reactions of linear, branched, and cyclic alkanes in the presence of NOx, Environmental Science & Technology, 43, 2328-2334, 2009.

Liu, Z., & Phillips, J. B.: Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, Journal of Chromatographic Science, 29, 227-231, 1991.

May, A. A., Presto, A. A., Hennigan, C. J., Nguyen, N. T., Gordon, T. D., & Robinson, A. L.: Gas-particle partitioning of primary organic aerosol emissions:(2) Diesel vehicles, Environmental Science & Technology, 47, 8288-8296, 2013.

McNutt, M. K., Chu, S., Lubchenco, J., Hunter, T., Dreyfus, G., Murawski, S. A., & Kennedy, D. M.: Applications of science and engineering to quantify and control the Deepwater Horizon oil spill, Proceedings of the National Academy of Science, 109, 20222-20228, 2012.

Phillips, J. B., & Xu, J.: Comprehensive multi-dimensional gas chromatography, Journal of Chromatography A, 703, 327-334, 1995.

671 Reddy, C. M., Arey, J. S., Seewald, J. S., Sylva, S. P., Lemkau, K. L., Nelson, R. K., Carmichael, C.
672 A., McIntyre, C. P., Fenwick, J., & Ventura, G. T.: Composition and fate of gas and oil released to
673 the water column during the Deepwater Horizon oil spill, Proceedings of the National Academy of
674 Science, 109, 20229-20234, 2012.
675
676 Reichenbach, S. E., Kottapalli, V., Ni, M., & Visvanathan, A. Computer language for identifying
677 chemicals with comprehensive two-dimensional gas chromatography and mass spectrometry,
678 Journal of Chromatography A, 1071, 263-269, 2005.
679 Riazi, M. R.: Characterization and properties of petroleum fractions. 1st Edition, American Society
680 for Testing Materials, MNL50, 2005
681
682 Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A.
683 P., Lane, T. E., Pierce, J. R., & Pandis, S. N.: Rethinking organic aerosols: semivolatile emissions
684 and photochemical aging, Science, 315, 1259-62, 2007.
685
686 Sakurai, H., Tobias, H. J., Park, K., Zarling, D., Docherty, K. S., Kittelson, D. B., McMurry, P. H.,
687 & Ziemann, P. J.: On-line measurements of diesel nanoparticle composition and volatility,
688 Atmospheric Environment, 37, 1199-1210, 2003.
689
690 Schauer, J. J., Kleeman, M. J., Cass, G. R., & Simoneit, B. R. T.: Measurement of Emissions from
691 Air Pollution Sources. 2. C1 through C30 Organic Compounds from Medium Duty Diesel Trucks,
692 Environmental Science & Technology, 33, 1578-1587, 1999.
693
694 Schauer, J. J., Kleeman, M. J., Cass, G. R., & Simoneit, B. R.: Measurement of emissions from air
695 pollution sources. 5. C1− C32 organic compounds from gasoline-powered motor vehicles,
696 Environmental Science & Technology, 36, 1169-1180, 2002.
697
698 Sonntag, D. B., Bailey, C. R., Fulper, C. R., & Baldauf, R. W.: Contribution of lubricating oil to
699 particulate matter emissions from light-duty gasoline vehicles in Kansas City, Environmental
700 Science & Technology, 46, 4191-4199, 2012.
701
702 Wang, F. C.-Y., Robbins, W. K., & Greaney, M. A.: Speciation of nitrogen-containing compounds in
703 diesel fuel by comprehensive two-dimensional gas chromatography, Journal of Separation Science,
704 27, 468-472, 2004.
705
706 Weitkamp, E. A., Sage, A. M., Pierce, J. R., Donahue, N. M., & Robinson, A. L.: Organic aerosol
707 formation from photochemical oxidation of diesel exhaust in a smog chamber, Environmental
708 Science & Technology, 41, 6969-6975, 2007.
709
710 Welthagen, W., Mitschke, S., Muhlberger, F., & Zimmermann, R.: One-dimensional and
711 comprehensive two-dimensional gas chromatography coupled to soft photo ionization time-of-flight
712 mass spectrometry: a two- and three-dimensional separation approach, Journal of Chromatography
713 A, 1150, 54-61, 2007.
714
715 Worton, D. R., Isaacman, G., Gentner, D. R., Dallmann, T. R., Chan, A. W., Ruehl, C., Kirchstetter,

716 T. W., Wilson, K. R., Harley, R. A., & Goldstein, A. H.: Lubricating oil dominates primary organic
717 aerosol emissions from motor vehicles, Environmental Science & Technology, 48, 3698-706, 2014.
718

719 Worton, D. R., Zhang, H., Isaacman-VanWertz, G., Chan, A. W., Wilson, K. R., & Goldstein, A. H.:
720 Comprehensive chemical characterization of hydrocarbons in NIST standard reference material
721 2779 Gulf of Mexico crude oil, Environmental Science & Technology, 49, 13130-13138, 2015.
722

723 Zielinska, B., Sagebiel, J., McDonald, J. D., Whitney, K., & Lawson, D. R.: Emission Rates and
724 Comparative Chemical Composition from Selected In-Use Diesel and Gasoline-Fueled Vehicles,
725 Journal of Air & Waste Management Association, 54, 1138-1150, 2004.
726
727

771 **Table 1.** Hydrocarbons identified in diesel fuel, lubricating oil and diesel emissions (gas and
772 particulate phases) with their respective m/z fragment ions and percentage mass contributions

Compound Class	m/z (M$^+$)*	% Mass closure		% Contribution to mass identified in emissions	
		Diesel Fuel	Lubricating Oil	Gas Phase	Particulate Phase
Total		**89.7**	**74.7**	**85.0 of TIC**	**75.0 of TIC**
$n + i$-Alkanes ($C_{11} - C_{33}$)	57 (C_nH_{2n+2})	62.2	23.0	39.9	47.3
Monocyclic Alkanes ($C_{11} - C_{33}$)	82 (C_nH_{2n})	13.8	35.6	17.4	19.6
Bicyclic Alkanes ($C_{11} - C_{33}$)	137 (C_nH_{2n-2})	5.0	9.2	9.7	7.5
Tricyclic Alkanes ($C_{11} - C_{33}$)	191 (C_nH_{2n-4})	<0.1	2.7	<0.1	2.7
Monocyclic Aromatics ($C_{11} - C_{33}$)	92, 119 (C_nH_{2n-6})	4.4	4.2	23.5	6.0
Bicyclic Aromatics ($C_{11} - C_{33}$)	128, 141	0.8	<0.1	2.0	
Adamantanes	135, 149, 163, 177				
Diamantanes	187, 188, 201, 215, 229				
Pentacyclic Alkanes	258, 272, 286				
Hexacyclic Alkanes	298, 312				
Steroids	239, 267				
Monoaromatic Steranes	253				
Steranes	217, 218				
Methyl Steranes	217, 218, 231, 232				
25-norhopanes	177				
28, 30-norhopanes	163, 191				
Hopanes	336				
PAHs		0.8		0.6	4.0
Biphenyls / Acenaphthenes		0.1		0.1	<0.1
Tetralin / Indanes	132, 145	2.5		6.9	
Oxygenates					7.1
Furanones	84				0.9
FAMEs	174				2.0
Miscellaneous Compounds					<3.0

*m/z ratios presented here are the main mass fragments present in the low ionisation energy mass spectra. CLIC expressions and 70eV mass spectra use more m/z fragments which were also used for qualification and quantification

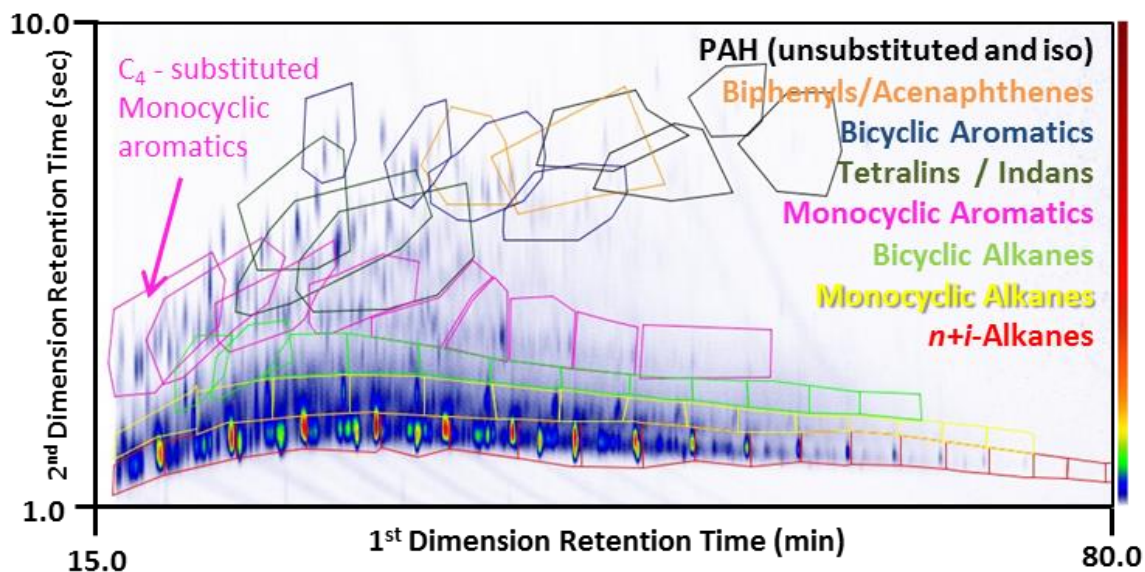



Figure 1. A contour plot (chromatogram with 70eV ionisation mass spectrometry) of diesel fuel separation. Peak height (intensity) increases with warmth (blue to red) of the colour scale. Each region fenced by a coloured polygon marks out the 2-dimensional chromatogram space in which are found isomers of a particular compound type having a particular carbon number (*e.g.* C₄–substituted monocyclic aromatics).

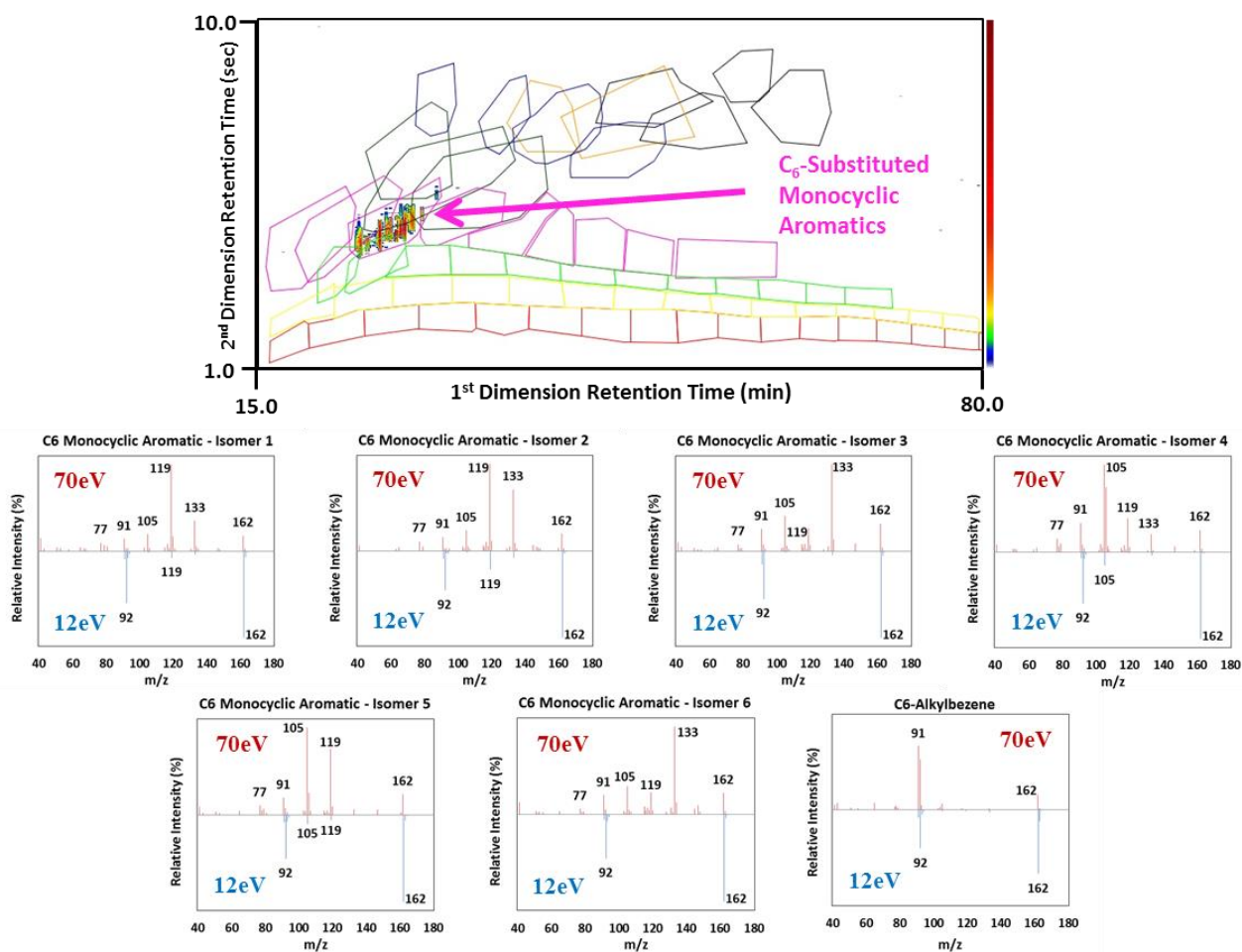



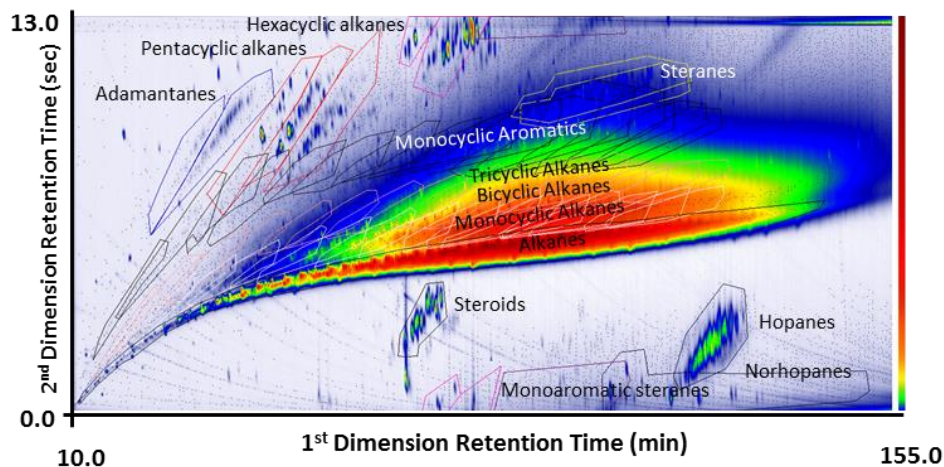
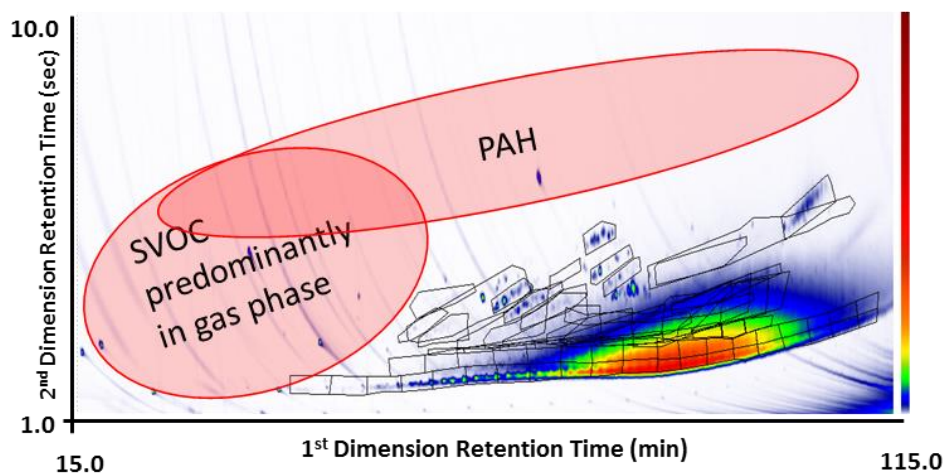
Figure 2. A contour plot (chromatogram with 12eV ionisation mass spectrometry) displaying C$_6$-substituted monocyclic aromatics identified by the CLIC expression. All C$_6$-substituted monocyclic aromatics are located within the pink polygon displayed. 70eV (red peaks) and 12eV (blue peaks) mass spectra corresponding to the peaks identified by the CLIC expressions in the SIC is shown for 6 different C$_6$-substituted monocyclic aromatics isomers.

(A)

(B)

Figure 3. A chromatogram (with 70eV ionisation mass spectrometry) of lubricating oil (5W30) **(A)** with labelled compound classes, using a methodology specific for characterising the composition of lubricating oil; **(B)** using methodology developed for characterising particulate phase emissions from diesel engine exhaust. Polygons drawn around sections of the chromatograms indicate compounds with different molecular masses within compound classes.

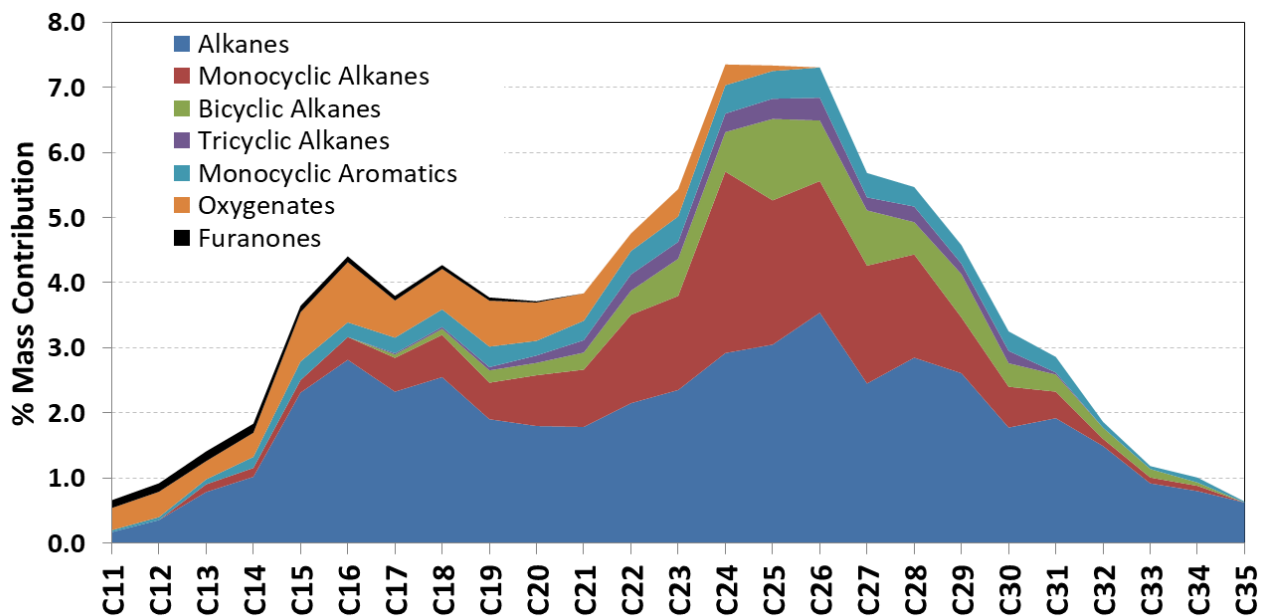



Figure 4. Percentage mass contribution of the compounds identified in homologous series as a function of carbon number in diesel exhaust particles.