

Characterisation of the chemical classes present in diesel fuel to identify the components that contribute to lubricity

by

Elize Smit

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae

In the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

Supervisor: Professor E.R. Rohwer

Co-supervisor: Doctor S. de Goede

April 2011

Declaration

I, Elize Smit, declare that the dissertation, which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature

Date

Acknowledgements

I would like to express my sincere gratitude to the following people and institutions:

My supervisor Prof. Egmont Rohwer as well as Drs. Stefan de Goede and Johan Coetzee for their guidance, support and motivation to successfully finish this study;

Sasol Technology, Fuels Technology for funding;

Dr. Peter Gorst-Allman, Yvette Naudé and Rina van der Westhuizen for all the practical advice and support;

Mr. David Masemula for the collection of liquid nitrogen;

My husband, Cornel Smit, for all the encouragement and patience;

My parents, Johan and Carla Hanekom, for their support and motivation;

My colleagues Belinda van der Westhuizen, Nina van Jaarsveld, Mia Ackermann, Shandré du Preez and Daniella Bezuidenhout for their support and camaraderie;

My family and friends for their understanding and support.

Elize

Contents

Summary	i
List of Abbreviations	ii
Chapter 1: Introduction	1
1. Preface	2
2. Objective	2
3. Experimental plan	3
4. Possible problems	4
5. Organisation	4
References	5
Chapter 2: Introduction to Petrochemistry and Tribology	6
1. Refining of oil	7
2. The Fischer Tropsch process	8
3. Tribology	10
References	16
Chapter 3: Gas Chromatography	18
1. Introduction	19
2. Revision of basic principles	20
3. Comprehensive two-dimensional gas chromatography	31
4. Properties of two-dimensional separation systems	39
5. Analysis of petrochemical samples	40
6. Conclusion: the importance of chemical characterisation	41
References	42
Chapter 4: Chemometrics	46
1. Chemometrics	47
2. What is a <i>model</i> ?	47

3. An overview of basic statistics	47
4. Linear regression	50
5. Linear association	53
6. Working with mixtures	55
References	59
 Chapter 5: Experimental Setup	 60
1. Samples	61
2. Lubricity measurement	61
3. Optimisation of GCxGC	62
4. GCxGC-TOFMS	68
5. Statistical analyses	85
References	87
 Chapter 6: GCxGC-TOFMS Results	 89
1. GCxGC-TOFMS results and data handling	90
2. Comparison of different column sets	92
3. Conclusion	109
References	110
 Chapter 7: Statistical Analysis of Results	 111
1. A look at the lubricity	112
1. Mixture models	113
2. Linear associations	116
3. Discussion	118
 Chapter 8: Conclusions and Final Remarks	 119
1. Main conclusions	120
2. Additional conclusions	122
3. Recommendations	122
 Appendix A: Two-dimensional chromatograms of samples	

Summary

Lubricity is the ability of a substance to prevent wear in a system where interacting surfaces are in close contact with each other. Some critical components in high pressure diesel injection systems are dependent on diesel fuel for lubrication and thus the lubricating properties of diesel fuel are very important. The relationship between the chemical composition and lubricating properties of diesel fuel were investigated, in an attempt to identify chemical classes and components in diesel that could improve lubricity. From literature, hetero-atom containing compounds and aromatic hydrocarbons were thought to be important.

The chemical compositions of diesel fuels were determined using comprehensive two-dimensional gas chromatography coupled to time of flight mass spectrometry. Four different column combinations were investigated and compared for the separation of these samples. The separations obtained for each column combination were compared graphically and mathematically. Better compound separation was achieved as the polarity of the stationary phase in the first dimension increased. Using a column with an ionic liquid stationary phase in the first dimension presented unique separations of aliphatic hydrocarbons but the improved compound separation resulted in the overlap of certain compound classes. The best class separation was achieved by using a polyethylene glycol column in the first dimension.

Multivariate statistical analysis was used to investigate the relationship between the chemical compositions and the measured lubricities of these fuels. A tentative statistical model was established, which showed that lubricity improved as the relative amount of aromatic hydrocarbons and oxygenates increased. The model also showed that oxygenates play a more important role than aromatic hydrocarbons in improving lubricity. The influence of chain length was also investigated and less volatile oxygenated compounds seem to contribute more to lubricity.

List of Abbreviations

1D	first dimension
2D	second dimension
1t_M	first dimension hold-up time
2t_M	second dimension hold-up time
1t_R	first dimension retention time
2t_R	second dimension retention time
α	selectivity
β	phase volume ratio
μ	linear flow velocity
\bar{v}	average linear rate of solute migration
σ^2	variance in peak profile
ℓ	length travelled in the column
Δt_R	difference in retention times for two analytes
w_{avg}	average peak width
r	correlation coefficient

AIHC	aliphatic hydrocarbons
ArHC	aromatic hydrocarbons
BOCLE	ball-on-cylinder lubricity evaluator
bp	boiling point
C_M	concentration of analyte in the mobile phase
CTL	coal-to-liquid
C_S	concentration of analyte in the stationary phase
d_f	film thickness
D_M	diffusion coefficient in the mobile phase
D_S	diffusion coefficient in the stationary phase
EI-MS	electron impact mass spectrometry
EIC	extracted ion chromatogram

FAME	fatty acid methyl ester
FID	flame ionisation detector
FT	Fischer Tropsch
GC	gas chromatography
GC-MS	gas chromatography – mass spectrometry
GCxGC	comprehensive two-dimensional gas chromatography
GC-GC	heart-cut two-dimensional gas chromatography
GCxGC-FID	comprehensive two-dimensional gas chromatography coupled to flame ionization detection
GCxGC-TOFMS	comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry
GLC	gas-liquid chromatography
GSC	gas-solid chromatography
GTL	gas-to-liquid
H_0	null hypothesis
H_A	alternative hypothesis
H / H_{tot}	theoretical plate height
HDI	hydrogen deficiency index
HFRR	high frequency reciprocating rig
HPLC	high performance liquid chromatography
HTFT	high temperature Fischer Tropsch
IL	ionic liquid
IR	infrared spectroscopy
k	retention factor
k_A	retention factor of compound A
k_B	retention factor of compound B
K_c	equilibrium / distribution constant
L	column length
LMCS	longitudinally modulating cryogenic system
LTFT	low temperature Fischer Tropsch
m/z	mass-to-charge ratio

M	mobile phase
m_M	mass of substance in the mobile phase
MS	mass spectrometry
m_s	mass of substance in the stationary phase
N	number of theoretical plates
NIST	National Institute of Standards and Technology
NMR	nuclear magnetic resonance
Ox	oxygenates
PDMS	polydimethyl siloxane
PEG	polyethylene glycol
R	resolution
S	stationary phase
SABS	South African Bureau of Standards
SANS	South African National Standards
SCD	sulphur chemiluminescence detectors
SFC	supercritical fluid chromatography
SL-BOCLE	scuffing load ball-on-cylinder lubricity evaluator
SRV	Optimol reciprocating rig
SSE	sum of squares of the residuals/errors
SSM	sum of squares attributable to the model
SST	sum of squares of total variability in the data
TIC	total ion chromatogram
TOFMS	time-of-flight mass spectrometry
V_M	volume of substance in the mobile phase
V_s	volume of substance in the stationary phase
WSD	wear scar diameter

Chapter 1: Introduction

Contents

1. Preface.....	2
2. Objective	2
3. Experimental plan	3
4. Possible problems.....	4
5. Organisation.....	4
References.....	5

1. Preface

The lubricating properties of diesel fuel are very important since older diesel injection systems rely exclusively on this fuel for lubrication. Some critical components in modern high pressure injection systems are also dependent on diesel fuel for lubrication. This research is aimed at investigating which chemical classes and components of a diesel fuel are responsible for the tribological behaviour of the fuel and the focus will be on lubricity.

The relationship between a fuel's lubricity and chemical composition has been studied previously [1-4]. Most studies were concerned with testing the lubricating properties of certain compounds by adding controlled amounts of these compounds to a fuel with low lubricity and empirically determining whether these compounds improved lubricity. The aim of this project is also to identify classes of compounds that may contribute to lubricity, but a different approach is taken. The lubricities of different diesel samples will be measured and compared to their chemical composition as determined by gas chromatography-mass spectrometry (GC-MS).

Multivariate statistical analysis will be used to find the relationship between the chemical class composition and the measured lubricity of a fuel. A similar approach to determine the relationship between the chemical composition of a sample and its physical properties has previously been described, although in this case nuclear magnetic resonance (NMR) was used as analysis method [5].

2. Objective

The objective of this project was to identify the components of diesel fuel responsible for improved lubricating properties. To achieve this, comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry (GCxGC-TOFMS) was to be used to determine different diesel sample compositions and compare these compositions with the lubricity of the samples (as measured by a high frequency reciprocating rig – HFRR), using a chemometric approach. Different

column combinations were to be investigated and compared to optimise the analysis of diesel fuel for this specific application. Since feed streams and final blending streams were to be analysed, the effect of some refining processes on lubrication properties would be better understood. The chemical interpretation of results will be a key factor in guiding the progress of the research.

3. Experimental plan

Diesel samples (unrefined streams and final blending products) will initially be analysed using GCxGC-TOFMS, from which chemical composition data will be obtained. Additional analysis methods may be employed later if sufficient information cannot be obtained from GCxGC-TOFMS analysis alone. Methods that may be additively used to complement GCxGC-TOFMS include NMR, infrared spectroscopy (IR) and supercritical fluid chromatography (SFC).

The lubricity properties of the samples (measured by HFRR) will be statistically compared with the composition data to identify the categories (grouped according to functional groups) of compounds that are related to lubricity. The different class(es) of compounds (grouped according to chain length and functional groups) present in the above categories will be considered to determine their relationship to lubricity. Once these compound class(es) are identified, the effect of individual components in these class(es) will be investigated. This approach is different to what has been carried out previously in literature [1-4].

The central objective is to understand why specific chemical compounds are more important than others when physical properties such as lubrication are considered. It will be important to look at the results from a chemical perspective and use chemical reasoning during interpretations.

4. Possible problems

The concentration of compounds that are deemed responsible for the lubricity of a sample could be so low that identifying and quantifying these peaks (using GCxGC-TOFMS) could be troublesome. During optimisation of the GCxGC-TOFMS system, this should be taken into account.

If the concentrations of above mentioned target compounds are so low that they cannot be distinguished from background noise (below the detection level of instrument), other methods may be needed to analyse these compounds as mentioned before. One method would be to increase the concentration of these compounds by selective removal of other compounds from the matrix. Another method would be to remove target compounds from the matrix and analyse them separately. This might involve extraction (e.g. methanol extraction), derivatisation or other reactions prior to analysis [6]. Alternative methods include separation using a silica column and SFC. The need for lower detection limits was to be evaluated in the initial stages of the research, although it was not originally anticipated.

5. Organisation

In Chapters 2, 3 and 4 literature reviews are given on petrochemistry and tribology, the chemical analysis of diesel and applicable chemometric methods. Chapter 5 explains the experimental setup used to achieve the aims of the research project. The optimisation of analytical techniques is also explained together with information on statistical methods and software used for chemometric analyses. The results obtained are presented and discussed in Chapters 6 and 7. The final conclusions and recommendations for future work are given in Chapter 8.

References

1. Wei, D.; Spikes, H.A. *Wear* **1986**, 111, pp. 217 – 235.
2. Anastopoulos, G.; Lois, E.; Zannikos, F.; Kalligeros, S.; Teas, C. *Tribology International* **2001**, 34, pp. 749 – 755.
3. Anastopoulos, G.; Lois, E.; Zannikos, F.; Kalligeros, S.; Teas, C. *Fuel* **2002**, 81, pp. 1017 – 1024.
4. Anastopoulos, G.; Lois, E.; Karonis, D.; Kalligeros, S.; Zannikos, F. *Energy* **2005**, 30, pp. 415 – 426.
5. Cookson, D.J.; Smith, B.E. *Energy & Fuels* **1990**, 4, pp. 152 – 156.
6. Link, D.D.; Baltrus, J.P.; Zandhuis, P.; Hreha, D.C. *Energy & Fuels* **2005**, 19, pp. 1693 – 1698.

Chapter 2: Introduction to Petrochemistry and Tribology

Contents

1.	Refining of oil.....	7
2.	The Fischer Tropsch process	8
2.1	Modes of operation.....	9
2.2	HTFT diesel vs. crude-derived diesel	10
3	Tribology.....	10
3.1	Lubricity regimes	10
3.1.1	Boundary lubrication	11
3.1.2	Fluid film lubrication	12
3.1.3	Mixed lubrication	12
3.2	Chemical nature of lubricants	12
3.3	Lubricity additives	13
3.4	Importance of lubricity in diesel	13
3.4.1	Ultra-low sulphur diesel fuels	14
3.4.2	Synthetic diesel fuels	14
3.5	Lubricity measurement	15
	References	16

1. Refining of oil

Crude oil is a fossil fuel that is formed from plants and animals that decompose underground over millions of years and it consists mainly of hydrocarbons [1]. Due to the chemical nature of crude oil, it is used as the conventional source of petrochemical fuels such as petrol, diesel and kerosene. To obtain these petrochemical products, crude oil is distilled into different boiling point fractions (see Figure 2.1). Although upgrading may be necessary, some of these fractions can be directly used as fuels whilst some fractions (e.g. heavier fractions) are refined to produce usable streams. These streams could be used to produce chemicals or blended to give fuels with desirable properties. Some of the very heavy fractions cannot be used as fuels or transformed into fuels cost-effectively and are thus further refined to produce lubricating oils, waxes and bitumen (also known as asphalt) [1]. Alternatively these fractions can also be cracked to give lower chain length molecules.

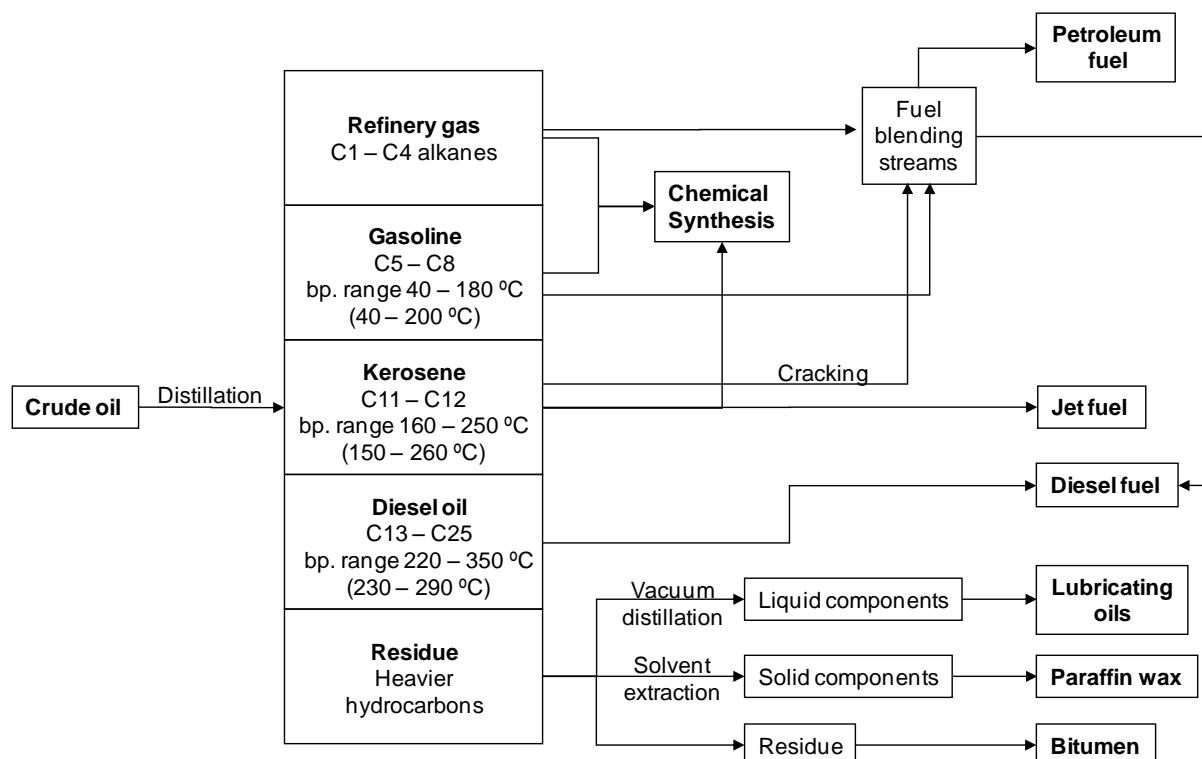


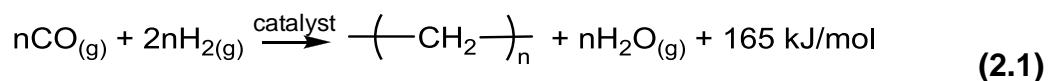
Figure 2.1 A schematic diagram of a typical crude oil refinery based on information obtained from [1]. Crude oil is distilled into boiling point (bp.) fractions, which are further refined to give fuels and chemicals as products. Values in brackets were obtained from [2] and converted from °F to °C.

For fuels to be sold commercially it is required that these fuels meet certain specifications. It may thus be necessary to add special chemicals known as additives, to a fuel in order to improve its quality. Such an additive is a compound or mixture of compounds that can alter a specific property of a fuel favourably with little or no undesirable effects on the other fuel properties (see Chapter 2, Section 3.3).

Crude oil sources are being depleted worldwide and the quality of the available crude oil is decreasing significantly [3]. For this reason, alternative fuels sources have been studied in detail over the last few years. These include coal, natural gas, biomass, hydrogen, methanol and many more [4]. In addition to being used directly as fuels, coal and natural gas can also be converted to synthetic fuels using the Fischer-Tropsch (FT) process.

2. The Fischer Tropsch process

German scientists Franz Fischer and Hans Tropsch developed the Fischer-Tropsch process in the 1920's [5]. It is in essence a process for converting carbon monoxide and hydrogen (commonly known as synthesis gas or syngas) to liquid hydrocarbons through catalysis. A simplified reaction equation of the overall reaction is given in Equation 2.1.



Coal and natural gas can be converted into liquid transportation fuels by the coal to liquid (CTL) and gas to liquid (GTL) processes. Sasol has been specialising in the production of synthetic fuels from coal and natural gas using the FT process since the 1950s [3].

In the CTL process, coal is gassified in the presence of water under high pressure and temperatures to produce synthesis gas. In the GTL process, natural gas is reformed to obtain an ideal ratio of carbon monoxide and hydrogen, which is also

used as syngas. In both processes the syngas is converted into hydrocarbons through the FT process (see Figure 2.2). Alternatively, syngas can be used to produce methanol.

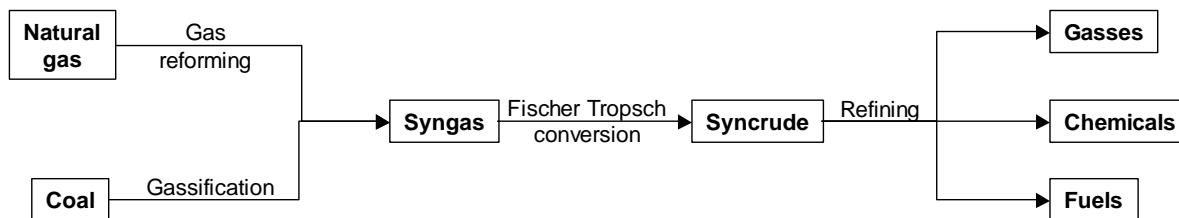


Figure 2.2 A simplified schematic diagram of a typical Fischer-Tropsch plant [7]. Coal or natural gas is used as starting material and converted into syngas. Syngas is transformed into syncrude and other products (hydrocarbon mixtures) via Fischer-Tropsch catalysis. The resulting streams are refined to produce fuel products, gasses and chemical products.

Different catalysts can be used for the FT process, including iron, cobalt, nickel and ruthenium [3,6]. Iron is most commonly used during high temperature FT (see Chapter 2, Section 2.1) since it is more cost-effective and easily obtainable. The nature of the catalyst will determine the final product slate produced by the process. Different mechanisms have been proposed in an effort to understand the final product composition better and to be able to predict what products can be obtained [6].

The final product streams from the FT reactor undergo separation (via distillation), upgrading and conversion to produce streams that can be blended together to give fuels with desirable properties. Some of the streams can also be used to produce certain chemicals. This refining process is in principle similar to that of natural crude oil although different conversion and upgrading steps may be used due to the different stream compositions [7].

2.1 Modes of operation

The FT conversion can be carried out at high temperatures (HTFT) or low temperatures (LTFT) [3,5,7]. From these two FT technologies different compounds are produced. The HTFT product streams tend to have more complex compositions especially in terms of heteroatom content [7]. The main LTFT product streams

typically contain long chain n-alkanes which can be used to produce different waxes. These streams can also be cracked to produce diesel blending streams. The product streams from HTFT are usually blended to obtain commercial fuels (such as diesel and petrol) [7].

2.2 *HTFT diesel vs. crude-derived diesel*

In general, synthetic fuels produced by HTFT contain less aromatic compounds compared to crude derived fuels. Synthetic fuels often contain olefins, which are only present in low concentrations in crude-derived fuels [3,7,8]. During HTFT conversion a large aqueous fraction is produced [3,7,8]. Some short chain oxygenates dissolve in this aqueous fraction known as reaction water [8]. It is thus clear that synthetic fuels need to be treated differently from crude-derived fuels with respect to refining and additives needed to comply with specifications.

3 Tribology

Tribology is the study of interacting surfaces in motion [9]. There are a few key physical properties of fluids, present between these interacting surfaces, which govern the way in which these surfaces will interact. Important tribological factors include viscosity, lubricity, wear and friction [9]. Although all the tribological properties of diesel fuels are important, this research will focus on fuel lubricity. Lubricity is a measure of the ability of a substance to prevent wear in a system where metal surfaces are in close contact with one another [9,10].

3.1 *Lubricity regimes*

Lubricity can be described in terms of three regimes. These regimes are boundary lubrication, fluid film lubrication and mixed lubrication as illustrated in Figure 2.3. To

distinguish between these regimes the thickness of the lubricating film or protective layer is considered with respect to the roughness of the surfaces [9,10]. Note that while gaseous, oil and solid lubrication do also exist, these phenomena are beyond the scope of this research.

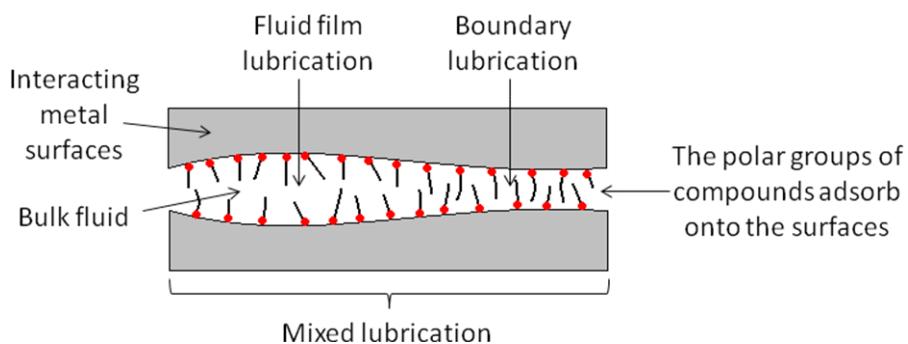


Figure 2.3 The three regimes of lubricity. The type of lubrication that takes place depends on the thickness of the lubricating film relative to the roughness of the interacting surfaces.

3.1.1 Boundary lubrication

A metal surface is usually covered with a metal oxide layer. It is important to remember that surfaces (irrespective of their appearance to the naked eye) are microscopically very rough; they contain hills (asperities) and valleys (grooves) [10]. When interacting surfaces are in very close contact, direct contact between the asperities can occur. If a monolayer of lubricating compounds is present on each of these interacting surfaces, boundary lubrication is observed [9,10].

Compounds such as fatty acids, consisting of polar and non-polar regions, can possibly prevent wear when interacting surfaces are in very close contact. The polar ends chemically or physically adsorb onto the metal oxide layer. These interactions are called chemisorption and physisorption respectively [9,10]. The non-polar part of the molecules will form a bristle-like structure above the surface of the metal. When two surfaces interact closely, these “bristles” will prevent the metal surfaces from contacting each other directly and thus prevent wear [9,10].

Note that it may only be possible for the hydrophobic regions to form “bristles” above the metal surface in the presence of a solvent. In its absence the hydrophobic

regions may lay flat on the surface. Thus the bulk of the fuel (alkanes) plays an indirect role in the lubricity. This is analogous to the way in which algae on a solid rock are only slippery when wet and swollen with water [11]. A similar phenomenon is observed in reversed-phase high-performance liquid chromatography (HPLC). The alkyl groups coupled to the surface of the silica packing material have to be solvated with an organic solvent prior to their efficient use as a stationary phase for the partitioning of analytes [12,13]. If the stationary phase is exposed to a highly aqueous mobile phase, the hydrocarbon chains fold over (aggregate) and thus retention of compounds is lost. This is also known as hydrophobic collapse [12,13].

3.1.2 Fluid film lubrication

When a relatively thick film of lubricant separates the interacting surfaces, it may result in fluid film lubrication [9,10]. This type of lubrication is known as hydrodynamic lubrication since the pressure within this film (that helps carry the load) is generated hydrodynamically (i.e. by the motion of the surfaces). The pressure can also be generated hydrostatically (from external pressure) leading to hydrostatic lubrication.

3.1.3 Mixed lubrication

Mixed lubrication occurs where hydrodynamic and boundary lubrication takes place simultaneously between interacting surfaces [9,10]. This may be caused by irregularities in the metal surface where close contact might be present in one region, whilst a fluid film can be formed in a neighbouring region.

3.2 *Chemical nature of lubricants*

The chemical compounds that contribute to lubricity usually have polar and non-polar regions, especially for boundary lubrication. Some compounds that have been studied for their lubrication behaviour include oxygenates (especially carboxylic

acids, alcohols, esters and ethers), nitrogen and sulphur containing compounds as well as aromatics [14-17].

3.3 Lubricity additives

Lubricity additives are frequently required (especially in severely hydrotreated fuels, see Chapter 2, Section 3.4.1) for sufficient lubrication. The chemical compositions of these diesel lubricity additives are important, but may differ from the compounds that occur naturally in diesel fuel which may improve lubricity.

Diesel lubricity additives adsorb onto the interacting surfaces to form monomolecular layers which effectively reduce the friction originating between the surfaces [18]. The most important compounds used as lubricity additives are C₁₂ to C₁₈ carboxylic acids, fatty alcohols as well as synthetic and natural esters of fatty acids [18]. The effect of these layer-forming agents increases with an increase in molecular mass. The layer-forming ability of the different polar groups increases in the order of alcohols, esters, unsaturated acids and saturated acids [18].

Fatty acid methyl esters (FAMEs) are the major constituents of first generation biodiesel. Due to the chemical nature of these compounds, biodiesel has been investigated as a possible lubricity additive for crude derived and synthetic fuels [19].

3.4 Importance of lubricity in diesel

Diesel engines are dependent on the fuel for fuel system lubrication and thus the lubricating properties of diesel fuel are very important. Diesel fuel lubricity is especially important for engines with rotary pump systems since no oil is present in the pump and thus the diesel is used for lubrication. It has been reported that fuels with inefficient lubricating properties may cause pump failure and blockages in the fuel system [20]. The lubricity of diesel fuel is also important for new generation

common rail engines where the pumps and injectors (between the pintle and seat) could fail due to insufficient lubricity.

Historically the importance of diesel lubricity was not always recognised since viscosity was seen as a measurement of wear protection provided by the fuel. It has since been noticed that viscosity and lubricity are not always directly related as assumed in the past. For instance, a fuel with high viscosity can have low lubricity whilst a fuel with low viscosity can have sufficient lubricity [14]. Even though lubricity and viscosity are not always directly related, some correlation does exist between these fuel properties.

As mentioned before, there are certain specifications that a fuel is required to meet before it can be sold commercially. Lubricity is one of the main requirements for a fuel to be fit for use. Today there are standardised methods to measure the lubricity of a fuel, which will be discussed later (see Chapter 2, Section 3.5). The necessity for these standards was first illustrated in Sweden, where several field problems were observed with the use of ultra-low sulphur diesel fuels [20].

3.4.1 Ultra-low sulphur diesel fuels

To decrease the amount of sulphur present in crude-derived oils (as governed by environmental concerns), fuels are subjected to hydrogenation (also known as hydrotreatment) whereby sulphur is lost via the formation of volatile hydrogen sulphide. Unfortunately the crucial components for lubricity such as carboxylic acids and other heteroatomic species (especially nitrogen containing compounds) are also hydrogenated and thus low sulphur diesel fuels tend to have very low lubricity [20,21]. To meet standard lubricity requirements, lubricity additives or biodiesel need to be added to the diesel (see Chapter 2, Section 3.3).

3.4.2 Synthetic diesel fuels

Even though sulphur-containing compounds do not necessarily contribute to lubricity, crude-derived fuels contain significantly larger amounts of these compounds. During

FT synthesis, no sulphur containing compounds are produced and the small amounts of sulphur containing compounds present in the final products originate only from the catalysts used [8]. As a result, hydrogenation is not required to remove sulphur-containing compounds. Synthetic fuels do however require hydrogenation to remove C₃ and C₄ carboxylic acids, which are known corroding agents [8]. Unfortunately longer chain carboxylic acids and other possible lubricating agents are also destroyed during this process, most likely contributing to the low lubricity of synthetic fuels.

3.5 Lubricity measurement

In South Africa the standard test method for measuring fuel lubricity is defined by South African National Standards (SANS), a division of the South African Bureau of Standards (SABS). The specifications and test method are based on ISO 12156-1 and CEC-F06-A-96 international specifications [22]. The HFRR is used according to the above specifications to measure lubricity and works as follows: A metal sphere is brought into close contact with a metal surface which is immersed in the sample to be analysed. The sphere is moved back and forth across the surface at a high frequency for a certain time, under controlled experimental conditions. The wear scar diameter (WSD) on the metal surface is measured using a microscope. A smaller WSD indicates better lubricity and vice versa. For a fuel to comply with standard South African specifications the corrected WSD should be no larger than 460 µm.

Other methods to measure the lubricity of a fuel have been developed including the ball-on-cylinder lubricity evaluator (BOCLE), the scuffing load ball-on-cylinder lubricity evaluator (SL-BOCLE) and the Optimol reciprocating rig (SRV) [23]. The methods for the measurement of lubricity differ mainly because they are based on the different lubricity regimes. Note that a similar study is currently being done at the Department of Chemical Engineering, University of Pretoria, on the comparison of these instruments for measuring lubricity and the different regimes in which they function.

References

1. *Oxford Dictionary of Chemistry*; Oxford University Press: Great Britain, **2004**.
2. McKetta, J.J. *Encyclopedia of chemical processing and design*; Marcel Dekker Inc.: United States of America, **1990**, Volume 13 pp. 238 - 245, Volume 35 p. 127.
3. Dry, M. *Catalysis Today* **2002**, 71, pp. 227 – 241.
4. Olah, G.A.; Goeppert, A.; Prakash, G.K.S. *Beyond oil and gas: the methanol economy*; Wiley-VCH: Federal Republic of Germany, **2006**.
5. Underwood, A.J.V. *Industrial & Engineering Chemistry* **1940**, 32, pp. 449 – 454.
6. Cornils, B.; Hermann, W.A. *Applied homogenous catalysis with organometallic compounds - Volume 2*; VCH: Federal Republic of Germany, **1996**, pp. 747 – 761.
7. Leckel, D. *Energy & Fuels* **2009**, 23, pp. 2342 – 2358.
8. De Klerk, A. *Catalysis Today* **2008**, 130, pp. 439 – 445.
9. Bhushan, B.; Gupta, B.K. *Handbook of Tribology*; McGraw-Hill Inc.: United States of America, **1991**, pp. 2.29 – 2.36
10. Hutchings, I.M. *Tribology: Friction and wear of engineering materials*; Edward Arnold (a division of Hodder & Stoughton Limited): Great Britain, **1992**, pp. 58 – 73.
11. Zhmud, B.; Roegiers, M. *Tribology & Lubrication Technology* **2009**, June, pp. 34 – 38.
12. Walter, T.H.; Iraneta, P.; Capparella, M. *Journal of Chromatography A* **2005**, 1075, pp. 177 – 183.
13. Yang, S.S.; Gilpin, R.K. *Journal of Chromatography* **1987**, 394, pp. 295 – 303.
14. Wei, D.; Spikes, H.A. *Wear* **1986**, 111, pp. 217 – 235.
15. Anastopoulos, G.; Lois, E.; Zannikos, F.; Kalligeros, S.; Teas, C. *Tribology International* **2001**, 34, pp. 749 – 755.
16. Anastopoulos, G.; Lois, E.; Zannikos, F.; Kalligeros, S.; Teas, C. *Fuel* **2002**, 81, pp. 1017 – 1024.

17. Anastopoulos, G.; Lois, E.; Karonis, D.; Kalligeros, S.; Zannikos, F. *Energy* **2005**, 30, pp. 415 – 426.
18. Mang, T.; Dresel, W. *Lubricants and Lubrication - second, completely revised and extended edition*; Wiley-VCH: Chichester, **2007**, pp. 107 – 113.
19. Wadumesthrige, K.; Ara, M.; Salley, S.O.; Ng, K.Y.S. *Energy & Fuels* **2009**, 23, pp. 2229 – 2234.
20. Tucker, R.F.; Stradling, R.J.; Wolveridge, P.E.; Rivers, K.J.; Ubbens, A. *SAE Paper* **1995**, 942016, pp. 62 – 76.
21. Spikes, H.A.; Wei, D.P. *Fuels International* **2000**, 1, pp. 45 – 65.
22. *South African National Standards (SANS) publication: Automotive diesel fuel—Edition 4*; **2006**, 342, p. 7.
23. Wielligh, A.; Burger, N.; Wilcocks, T. *Industrial Lubrication and Tribology* **2003**, 55, pp. 65 – 75.

Chapter 3: Gas Chromatography

Contents

1.	Introduction	19
2.	Revision of basic chromatography principles	20
2.1	Resolution	20
2.2	Temperature programming.....	24
2.3	Important chemical interactions.....	25
2.4	Linear and non-linear chromatography.....	27
2.5	Peak capacity	30
3.	Comprehensive two-dimensional gas chromatography.....	31
3.1	Instrument	31
3.1.1	Modulators for GCxGC.....	33
3.1.2	Detectors for GCxGC	34
3.2	Data analysis	36
3.2.1	Classification of compounds into chemical groups.....	38
4.	Properties of two-dimensional separation systems	39
4.1	Enhanced peak capacity	39
4.2	Orthogonality	40
5.	Analysis of petrochemical samples	40
6.	Conclusion: the importance of chemical characterisation.....	41
	References	42

1. Introduction

Chromatography is a powerful analytical methodology during which a mixture of analytes is separated into fractions based on their different physical and chemical properties [1]. The chemical nature of an analyte will govern its partitioning between the mobile (M) and stationary phases (S). The stronger the intermolecular forces of attraction between the analyte and the stationary phase, the greater the partitioning into the S. This will cause the compound to be more retained than an analyte which experiences weaker interactions and consequently these compounds can be separated.

In gas chromatography (GC) hydrogen, helium or nitrogen gas is used as the M (i.e. the carrier gas). The liquid S is usually a polymer of siloxane with substituents that vary in chemical structure, e.g. polarity, aromaticity and chirality. Other stationary phases that are not based on siloxanes also exist, e.g. polyethyleneglycol (PEG) and ionic liquid (IL). The liquid stationary phase can be used to coat the walls of the glass capillary column, or be cross-linked and covalently bonded to the walls [2].

In comprehensive two-dimensional gas chromatography (GCxGC), two capillary columns are attached in series. A mixture of analytes is separated into fractions in the first column based on a specific physical or chemical property. These fractions are captured and reinjected (modulated) onto the second column where they are separated into sub-fractions based on a different property [3]. These two separations are known as the two “dimensions” of GCxGC, e.g. if the first column has a polydimethylsiloxane (PDMS) stationary phase, compounds are separated based on volatility or vapour pressure in the first dimension (^1D). If the second column contains a PEG stationary phase, compounds are separated based on polarity in the second dimension (^2D).

2. Revision of basic chromatography principles

The behaviour of a chemical compound during a GC analysis is determined by its chemical nature and the environment within the capillary column. These factors are illustrated in Figure 3.1 and will be considered in more detail later. Different analytes have different properties and thus will behave differently in similar environments [1]. This principle is exploited during GC in order to separate a mixture of analytes. Pure compounds elute from the GC system as symmetric peaks and (ideally) have a Gaussian (normal) distribution [4]. The apex of each peak indicates when most of the molecules elute and this is known as the retention time [4]. The area under each peak can be determined to give an idea of the amount of this compound present in the mixture.

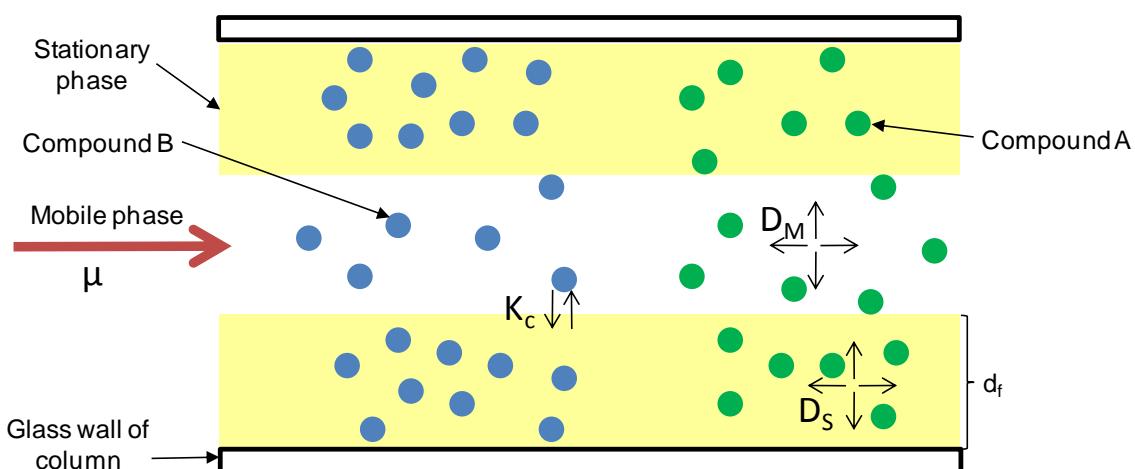


Figure 3.1 Factors influencing the behaviour of different chemical compounds during GC. The type of stationary and mobile phase influences the separation between compounds A and B. The thickness of S (d_f) is also important. The equilibrium constant (K_c) describes the concentration partitioning between M and S. Other factors such as the linear velocity of the mobile phase (μ) and diffusion in M and S (described by the respective diffusion coefficients; D_M and D_S) also influence the separation.

2.1 Resolution

Since the main objective of GC (and other forms of chromatography) is to separate a mixture of analytes, the degree of separation achieved can be used to compare different systems. This separation can be quantified in terms of resolution [5]. To

determine the resolution between two peaks, their respective retention times and peak widths need to be considered [5]. The plate theory [6,7] or the newer rate theory [8] can be used to explain retention times and peak widths. Rate theory is however far superior in explaining the variation in peak widths with factors such as linear flow rate of M, diffusion coefficients and column dimensions (internal diameter and film thickness).

Chromatographic plate theory [6,7] is based on an equivalent description of the distillation process and can be used to emphasize the relationship between distillation (separation based on boiling point) and chromatography (separation based on chemical interactions). In a distillation column, there are a specified number of plates (N). The plate height (H) can be defined as the distance between two consecutive plates and calculated by dividing the column length (L) by N . At each of these plates the analytes are in perfect equilibrium between the gas and liquid phase [6]. This is similar to gas chromatography where the analytes partition between the mobile (gas) phase and stationary (liquid) phase. At the apex of a peak the compounds are essentially in equilibrium between M and S.

The equilibrium (or distribution) constant (K_c) is used to describe this partitioning, i.e. the concentration of analyte in S (C_S) relative to the concentration in M (C_M) [9]. Concentration can be described as the amount of substance in a specified volume, and likewise K_c consists of two parts; the phase volume ratio (β) (volume of substance present in M relative to S; V_M and V_S respectively) and the retention factor (k) (mass of substance present in S relative to M; m_S and m_M respectively) as shown in Equation 3.1 [9].

$$K_c = k \times \beta \quad \text{where } k = m_S/m_M \text{ and } \beta = V_M/V_S \quad (3.1)$$

The *retention factor* (k) and the linear velocity of the mobile phase (μ) can be used to describe the average linear rate of solute migration (v) as shown in Equation 3.2 [10]. Different compounds have different K_c values and consequently they are retained more or less and elute at different times (different v). Analytes that spend

more time in M will elute earlier than analytes that are more retained (spend more time in S).

$$\bar{v} = \mu \times \frac{1}{1+k} \quad (3.2)$$

During distillation, the separation of analytes can be improved by increasing the number of plates. This can be done by increasing the column length and/or decreasing the plate height [6]. In GC the “number of plates” can also be increased by increasing the column length and/or decreasing the *theoretical* plate height (there are no physical plates as in distillation). By doing this, sufficient separation can be obtained even when the K_c 's of the different analytes are only slightly different.

The rate theory emphasizes the non-equilibrium nature of a moving profile in the column, thus the kinetics present in the separation system, and can best be used to explain the variation in peak width [8]. Narrow peaks are easier to resolve from one another, and consequently peak broadening should be minimised. As described above, compounds are in equilibrium between M and S. The rate at which analytes partition between these phases (C) will determine how much the peak will broaden. Other factors that also influence peak broadening are longitudinal diffusion in the M (B), the presence of multiple flow paths (A) and μ as shown in the Van Deemter equation (Equation 3.3) [8].

$$H_{\text{tot}} = A + B/\mu + C\mu \quad (3.3)$$

H_{tot} is the theoretical plate height, which is now defined as the variance in peak profile (σ^2) introduced per unit length moved in the column (ℓ) (Equation 3.4) [7]. The linear velocity (μ) is usually reported as centimetres travelled per second and is related to the flow rate (millilitres per minute) by the cross-sectional area of the column [8].

$$H_{\text{tot}} = \sigma^2/\ell \quad (3.4)$$

A, B and C represent the different mechanisms that contribute to peak broadening [8]. For open tubular columns, multiple flow paths (A) are not present (compare packed columns) and thus cannot influence peak broadening. As a result A = 0 [8]. B explains the peak broadening due to longitudinal diffusion (D_M), which is inversely proportional to μ . When μ is increased, compounds spend less time in the system and thus will have less time to diffuse and consequently less peak broadening is observed [8]. C explains the peak broadening due to the resistance to mass transfer which can be described as the limited speed with which the equilibrium between analyte in M and S can be re-introduced once disturbed by the moving mobile phase. Thus the rate of partitioning between M and S is crucial. If an analyte reaches the equilibrium (between M and S) very quickly, less peak broadening will be observed [8]. So the determining factor here is how easily a molecule will be able to move into and out of the stationary phase. This is determined by the diffusion in the stationary phase (D_S), d_f , μ and k .

The above factors which influence the movement of analytes through a chromatographic system can further be related to the separation achieved between two or more compounds. The degree of separation achieved for two pure compounds can be quantified in terms of resolution (R) and depends on the difference in retention times between the two analytes (Δt_R) and their average peak width (W_{avg}) as shown in Equation 3.5 [5].

$$R = \Delta t_R / W_{avg} \quad \text{where } W_{avg} = 4\sigma \quad (3.5)$$

Consider a situation where compound B is more retained compared to compound A (see Figure 3.1), the relative retention of these compounds is known as the *selectivity* (α) as shown in Equation 3.6, where k_A and k_B are the retention factors of compounds A and B respectively [11].

$$\alpha = k_B / k_A \quad (3.6)$$

Taking into account above statements, the resolution in a chromatographic system can be shown to depend on its *efficiency*, *selectivity* and *retention* capability

(Equation 3.7). To get the best possible resolution between two compounds, all these factors need to be maximised [5].

$$R = \frac{\sqrt{N}}{4} \times \underbrace{\frac{(\alpha - 1)}{\alpha}}_{\text{selectivity}} \times \underbrace{\frac{k}{1 + k}}_{\text{retention}}$$

Factors based on: efficiency selectivity retention (3.7)

The number of theoretical plates (N) is a measure of the *efficiency* of the system [7]. Note that N and H are related to one another by the column length as discussed previously. The Van Deemter equation (Equation 3.3) can be used to minimise H , and thus maximise the *efficiency* of the system.

The *relative retention* or *selectivity* (Equation 3.6 and 3.7) depends on the chemical nature of the stationary phase and the chemical nature of the two analytes to be separated [11]. If an analyte is chemically similar to the stationary phase, it will be retained more than an analyte with dissimilar chemical characteristics resulting in a larger relative retention and α . This is based on the fact that “like dissolves like” and consequently it is preferred to select a column with a chemical nature corresponding to that of the analyte for which extra retention is desired [12].

The last factor regarding *absolute retention* (Equation 3.7) approaches 1 as k increases. It is suggested to work with a k -value between 2 and 5 [13], since larger k -values will result in long analysis times with small gain in resolution. It is important to note that k is dependent on temperature, and the above statements (as well as Equation 3.7) are only applicable to isothermal conditions. It is possible to calculate the *absolute retention* of an analyte for non-isothermal conditions, but it is not as straight-forward to interpret as the isothermal equation [14].

2.2 Temperature programming

When two compounds co-elute they can be resolved by retaining one compound more than the second compound (affecting their relative migration rates) or by

decreasing the peak width. In GC more than two peaks usually have to be resolved and the general elution problem is encountered [15]. This is where a lower temperature is required in order to resolve compounds that elute early in the elution profile (too small k -values). The more retained compounds will thus take much longer to elute. To circumvent this, temperature programming is introduced where the temperature is gradually increased [15]. Thus early eluting compounds can be separated (at lower temperatures to provide sufficient k -values) first. The more retained compounds can then be caused to elute faster with higher temperatures (which provides for lower k -values). When using temperature programming, the resolution achieved remains the same as in the isothermal case, but the analysis time is significantly reduced (the k -values of compounds vary over a wide range under isothermal conditions).

2.3 *Important chemical interactions*

Depending on the chemical nature of the stationary phase, several different intermolecular interactions with varying strengths are possible between the analytes and chemical groups in the stationary phase [16]. These interactions govern the way in which partitioning of analytes between M and S occurs. The chemical nature of the interacting groups will determine the type of intermolecular interactions as shown in Table 3.1.

The relative strength of the interactions decreases as the polarity or polarisability of both interacting molecules decreases. When ionic compounds interact with other (polar or polarisable) molecules the attraction forces are much stronger compared to other intermolecular forces. Hydrogen bonds can also be seen as dipole-dipole interactions, but the dipole moments of the interacting molecules are so large that a considerably stronger attraction force is present. The Van der Waals forces indicated in Table 3.1 are the weakest intermolecular interactions compared to the other interactions shown. London dispersion forces are the weakest of all the Van der Waals forces [16].

Table 3.1 Intermolecular interactions that can occur between the stationary phase and analytes during a GC separation.

Column	Stationary phase		Analyte		Intermolecular interaction
	Chemical nature	Example	Chemical nature	Example	
Highly polar ¹	Ion	IL	Permanent dipole	CHCl ₃	Ion – dipole
	Ion ²	IL	Delocalised π-electrons ³	C ₆ H ₆	Ion – induced dipole (Ion –π interaction)
Polar	H bonded to N/O/F	PEG	H bonded to N/O/F	R-OH	Hydrogen bond
	Permanent dipole	PEG	Permanent dipole	CHCl ₃	Dipole – dipole ⁴
	Permanent dipole	PEG	Delocalised π-electrons ³	C ₆ H ₆	Dipole – induced dipole (Debye forces) ⁴
	Permanent dipole	PEG	Induced dipole	C ₆ H ₁₄	Dipole – induced dipole (Debye forces) ⁴
Medium polar	Delocalised π-electrons ³	50% phenyl / 50% PDMS	Permanent dipole	CHCl ₃	Dipole – induced dipole (Debye forces) ⁴
	Delocalised π-electrons ³	50% phenyl / 50% PDMS	Delocalised π-electrons ³	C ₆ H ₆	π-π interactions (π stacking)
Non-polar	Induced dipole	PDMS (only 1% or 5% phenyl)	Permanent dipole	CHCl ₃	Dipole – induced dipole (Debye forces) ⁴
	Induced dipole	PDMS (only 1% or 5% phenyl)	Induced dipole	C ₆ H ₁₄	Induced dipole – induced dipole (London dispersion forces) ⁴

¹ Specifically applicable to the SLB-IL100 column, other ionic liquids used as stationary phases may be less polar than conventional stationary phases

² When an imidazolium group is present (contains a conjugated unsaturated system), π-stacking interactions can also occur with compounds containing delocalised electrons

³ Conjugated unsaturated systems and aromatic compounds

⁴ Van der Waals forces

Different columns can be compared based on their stationary phases and the types of interactions (which govern compound retention) that are possible. A column containing a S with polar or polarisable functional groups is generally known as a

polar column, e.g. a column with a PEG stationary phase is known as a polar column compared to a column with a 5% phenyl / 95% PDMS stationary phase. A column with a 50% phenyl / 50% PDMS stationary phase is said to have medium polarity. The polarity of columns containing IL stationary phases will depend on the nature of the cation and anion of the IL. The SLB-IL100 column (see Chapter 5, Section 4.2) can be described as highly polar.

2.4 Linear and non-linear chromatography

An important factor in chromatography is the peak shapes obtained for certain compounds. It is always preferred to work under linear chromatographic conditions, i.e. where the ratio C_s/C_m (i.e. K_c) is constant (independent of total concentration). In this ideal situation, peaks with Gaussian distributions are obtained [4]. Under non-linear conditions, this concentration ratio (K_c) is altered (even within a peak as analyte concentration changes). This is due to certain interactions that result in peak tailing and fronting [17]. This often occurs when the immediate environment in the capillary column changes.

It is important to differentiate between a profile in space and a profile in time. Movement in space results in a length profile of concentration inside the column at a specific time (snapshot). Movement in time can be measured as a concentration profile exits the column (at a fixed length) as seen in the time-based chromatogram. Less retained molecules reach the detector earlier and thus the concentration profile in time is a mirror image of the concentration profile in space.

Thus far only gas chromatographic systems with a liquid stationary phase (gas-liquid chromatography, GLC) have been considered. In a system where the stationary phase is a porous solid (gas-solid chromatography, GSC), peak tailing can be described in terms of the Langmuir adsorption isotherm (Figure 3.2) [17]. In this situation K_c becomes smaller with increased analyte concentration. The porous S can accommodate a large amount of molecules due to its large surface area. At lower analyte concentrations K_c stays constant because the porous S can

accommodate all the analyte molecules that will tend to partition into S. When the amount that can be accommodated is exceeded, the remaining molecules will stay in M (they cannot partition into S). This occurs at higher analyte concentrations and C_M is larger than C_S . The molecules in M move forward (due to the constant flow rate of M), whilst the molecules in S lag behind [17] as shown in Figure 3.2. Peak tailing is consequently observed in the chromatogram (movement in time).

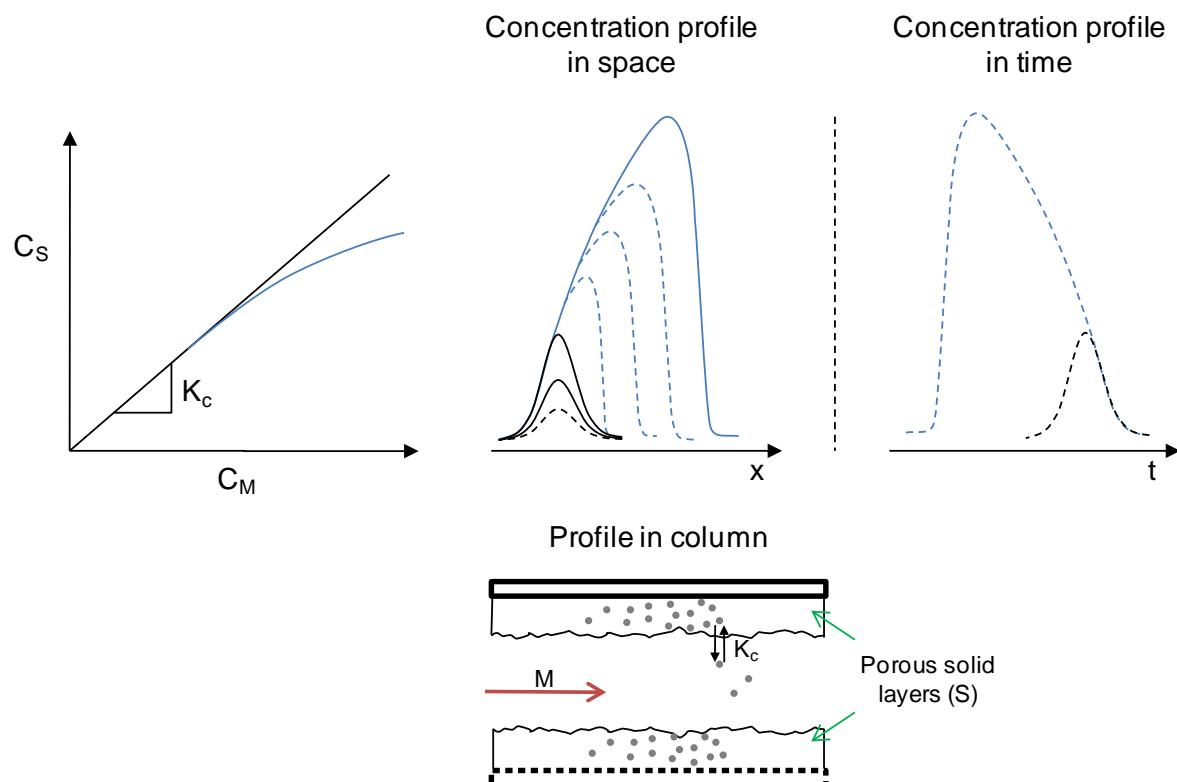


Figure 3.2 The Langmuir isotherm describing the situation where K_c becomes smaller with an increased analyte concentration (shown in blue) in GSC. The movement in space and time is also shown.

The anti-Langmuir adsorption isotherm describes the situation where K_c becomes larger at increased analyte concentrations [17] as shown in Figure 3.3. At high concentrations the stationary phase (in this case a thin film of liquid, GLC) becomes saturated, and interactions among analyte molecules are favoured above interactions with S (this occurs more readily as the chemical nature of the analyte and S differ more significantly) [17]. Consequently C_S becomes larger than C_M , and the result of this is peak fronting as shown in Figure 3.3. At lower analyte concentrations the interactions among analyte molecules are limited due to the limited amount of molecules present and thus peak fronting will occur to a lesser extent.

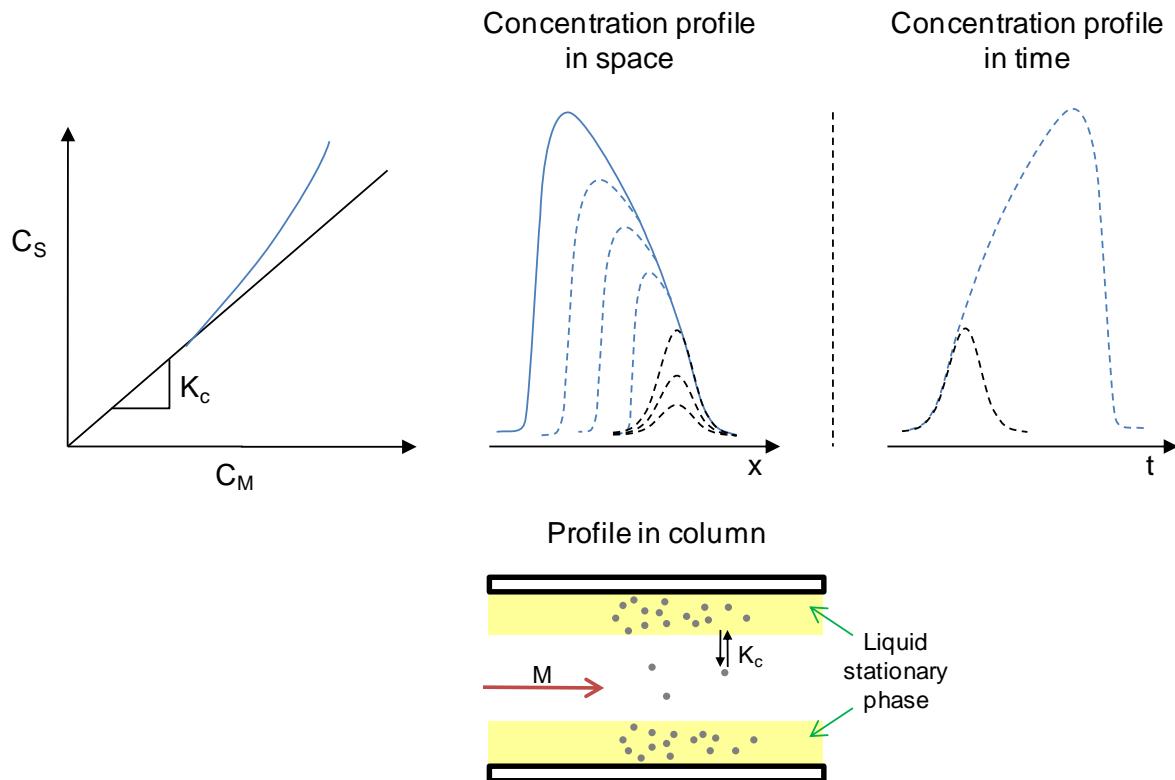


Figure 3.3 The anti-Langmuir isotherm describing the situation where K_c becomes larger with an increased analyte concentration (shown in blue) in GLC. The movement in space and time is also shown.

Peak skewing (caused by Langmuir or anti-Langmuir behaviour) can be avoided by diluting the sample or injecting a smaller amount of the sample into the column [18]. This is especially true when using columns designed for fast analyses, where the internal diameter is small and the film of stationary phase is thin (these columns have a small capacity).

In GLC it is possible to observe peak tailing (Langmuir behaviour), but due to a different mechanism [17]. The silanol groups of the capillary wall are usually deactivated before it is coated with the stationary phase. As the column ages or in cases where the capillary wall was not sufficiently deactivated, reactive groups could be accessible. If a polar (or polarisable) analyte molecule encounters such an active site by chance it will interact very strongly with these groups and result in peak tailing [17]. This type of peak tailing can be observed at very low concentrations (compared to peak tailing discussed above). At very high concentrations, the peak tailing becomes less significant since the limited number of active sites become saturated [17].

A single peak of e.g. a specific fatty acid contains both high and low concentration regions in the peak profile. In the regions where low concentrations are present, Langmuir behaviour is observed due to interactions with exposed silanol groups, which results in peak tailing. In the regions where high concentrations are present anti-Langmuir behaviour is observed due to overloading, which results in peak fronting. Thus for a single peak, both peak fronting and tailing can occur simultaneously.

Polar compounds such as carboxylic acids and alcohols tend to tail when a less polar column is used to separate these compounds. When a very low concentration of an analyte is present and significant tailing occurs, it might be difficult to recognise the peak, due to its very large peak width and short peak height. If a more polar column (e.g. PEG) is used to separate polar compounds, peak tailing is not observed due to the preferential adsorption of molecules onto the end-groups (alcohol moieties) of the stationary phase.

2.5 Peak capacity

For very complex mixtures, resolution of all compounds is seldomly attainable. This is due to the infinite resolving power of the chromatographic system. For each system there is a maximum number of ideally spaced peaks that can be resolved under optimum conditions. This is known as the peak capacity, a concept introduced by Giddings [19]. Thus if the number of compounds (in the mixture to be analysed) is more than the peak capacity, co-elution will inevitably take place. This is even more evident if peaks are not ideally spaced but randomly spaced in the chromatogram [20].

If one considers a one-dimensional GC system where a complex mixture is separated based on e.g. volatility, compounds with similar vapour pressures will co-elute even under optimum conditions (as discussed above). To resolve such peaks, longer columns with thinner stationary films and smaller internal diameters are

necessary, together with longer analysis times. This is not very practical, and an alternative approach is needed.

3. Comprehensive two-dimensional gas chromatography

GCxGC is ideal for the analysis of very complex samples [3,21]. During a GCxGC analysis, a sample is separated into different fractions (similar to one dimensional gas chromatography) and then exposed to a second separation (based on a different chemical property than the first) [3]. The technique is called “comprehensive” because the full sample is subjected to both separations, contrary to heart-cut two-dimensional gas chromatography (GC-GC) where only certain fractions are collected for further separation [22].

Since compounds are separated based on two different chemical properties, it is effectively separated in two different systems. The peak capacity of a comprehensive system, such as GCxGC, is calculated as the product of the peak capacities of the two different systems. This makes GCxGC a powerful separation method in which the number of compounds that can be resolved increases significantly compared to a one dimensional system (see Chapter 3, Section 4.1).

3.1 *Instrument*

A schematic diagram of a GCxGC system is given in Figure 3.4. The first and second columns (Figure 3.4, nr. 2 and 6 respectively) are housed in separate ovens (Figure 3.4, nr. 3 and 7 respectively) which can be controlled individually. The two columns are connected with a “pressfit connector” (Figure 3.4, nr. 4) and modulation occurs on the second column between this connection and the secondary oven. Fractions are captured and reinjected onto the second column by the modulator (Figure 3.4, nr. 5). The second column is connected to a detector (Figure 3.4, nr. 8).

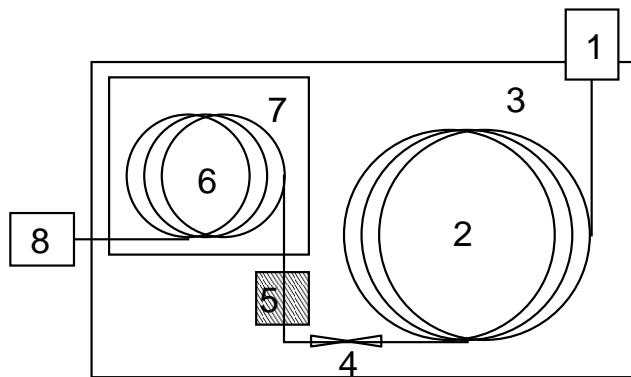


Figure 3.4 A schematic diagram of a GCxGC instrument. The instrument consists of an inlet (1), a primary column (2), a primary oven (3), a modulator (5), a secondary column (6), a secondary oven (7) and a detector (8). The two columns are connected with a pressfit connector (4).

The first column is usually a long column (30 to 60 m) with a medium bore width (0.2 to 0.3 mm). The film thickness of this column ranges between 0.2 and 0.3 µm. Under optimum conditions, peaks with widths of 10 to 20 seconds are usually obtained from such a column. This is mainly due to the low temperature programming rate in the first dimension, which is usually between 1 and 3 °C/min [21]. Narrower peaks in time can be obtained by using faster programming rates even though resolution will not be improved by doing this. This is not ideal for GCxGC, where the slower elution time of peaks in the first dimension is crucial for efficient modulation (the way in which the analytes are transported from one separation to the next) and to allow sufficient time for the second separation.

Every peak that elutes from the first dimension needs to be modulated at least 3 to 4 times in order to retain the separation achieved in the first dimension [23]. Other literature references indicate that more modulations per peak (4 to 5) are necessary [21]. More modulations per peak are preferred in order to preserve the peak shape in the first dimension. This is important since the shape of a chromatographic peak contains a lot of information.

For a ${}^1\text{D}$ peak which is 15 seconds wide, a part of it needs to be captured and reinjected onto the second column every five seconds for it to be modulated three times. This means that the analysis time of the second dimension cannot exceed five seconds. This is also known as the modulation period [22].

To achieve sufficient separation in e.g. five seconds, the second column has to be very short (1 to 2 m) and narrow (0.1 to 0.2 mm) with a very thin film of stationary phase (0.1 to 0.2 µm). Under optimum conditions, peaks with widths between 100 and 600 milliseconds are obtained [21]. The ²D separation occurs in such a short time, that it can be considered to be isothermal (the temperature usually increases with 2 °C every minute).

3.1.1 Modulators for GCxGC

The modulator (Figure 3.4, nr. 5) is the heart of the instrument since it governs the way in which fractions are transported from one dimension to the next. Since the development of GCxGC as a separation technique, several different modulators have been developed. These include thermal modulators, e.g. a slotted heater (“sweeper”) [24] and a longitudinally modulating cryogenic system (LMCS) [25] and valve-based modulators. The dual-stage jet modulator [21,26] described below is most commonly used today.

The dual-stage jet modulator consists of two hot air jets and two cold N₂ (g) jets [26]. These jets are used to capture and reinject (modulate) eluting substances into the secondary column as illustrated in Figure 3.5.

N₂ (g) is cooled as it passes through a coil immersed in liquid nitrogen on the way to the modulator. Pressurised air cannot be used for the cold jets, since the lower vapour pressure oxygen in the air will condense and cause blockages in the system. The first cold jet (Figure 3.5, nr. 1) traps the compounds eluting from the first column onto the first part of the second column. The first hot air jet (Figure 3.5, nr. 2) is then turned on to release this trapped fraction [26].

The second cold jet (Figure 3.5, nr. 3) captures and refocuses the fraction, which is then reinjected with the second hot air jet (Figure 3.5, nr. 4). While the fraction is being refocused and reinjected in the second part, the first cold jet captures the next fraction eluting from the first dimension [26]. This cycle is repeated throughout the entire run every five seconds if the modulation period is set at five seconds.

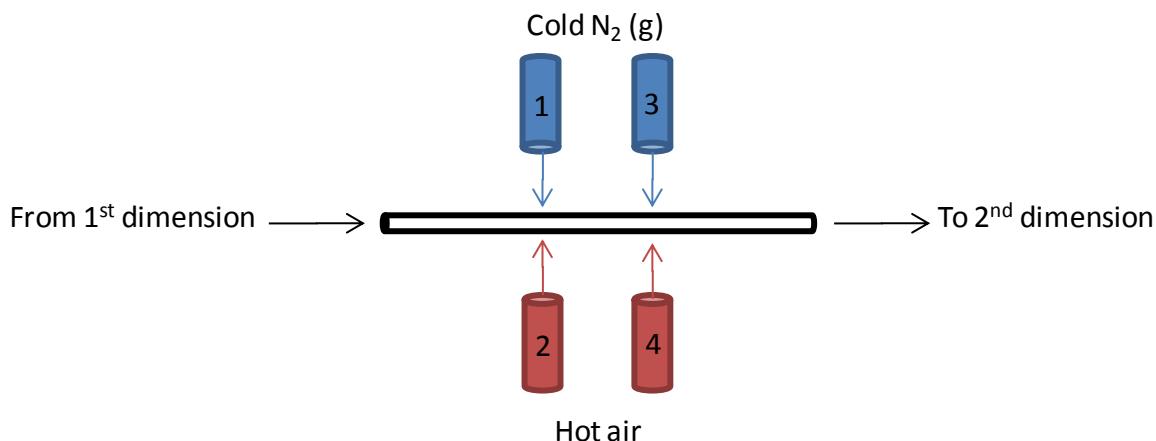


Figure 3.5 A schematic diagram of a dual-stage jet modulator. The first cold jet (1) traps the compounds eluting from the first column. The first hot air jet (2) then releases this trapped fraction. The second cold jet (3) captures and refocuses the fraction, which is then reinjected with the second hot air jet (4). While the fraction is being refocused and reinjected in the second part, the first cold jet captures the next fraction eluting from the first dimension [26].

3.1.2 Detectors for GCxGC

Several detectors can be used in conjunction with GCxGC (Figure 3.4, nr. 8). It is important that the detector should have a small internal volume and a fast response time and acquisition rate (20 – 100 Hz), since the compounds elute rapidly from the second column [21].

Examples of detectors that have been used include flame ionisation detection (FID) [21], mass spectrometry (MS) including quadrupole [27] and time-of-flight (TOFMS) analysers [28] and sulphur chemiluminescence detectors (SCD) [29]. Of these detection methods, FID and TOFMS are most frequently used in research due to their high speed and the wide applicability of these analysers as detectors for GCxGC.

GCxGC-FID is the preferred analysis method for quantification of petrochemical samples since the (mass) response factors for FID are similar for different hydrocarbons and more repeatable than those for electron-impact MS (EI-MS) [30]. The response factors for EI-MS are governed by the probability of total ionisation and thus the total response is not only dependent on the amount of substance, but also the ionisation efficiency of the substance. Note that even though compositional data cannot be obtained from GCxGC-TOFMS without the use of internal standards,

the relative compositions of several samples can be compared, given that the same MS parameters are used during analysis. These results are however still not comparable to GCxGC-FID results.

The TOFMS is an ideal mass selective detector for GCxGC analyses, since it has a fast acquisition rate of up to 500 spectra per second [31]. A full mass spectrum is obtained for each peak compared to what would be possible with quadrupole MS (which is limited by a slow scan rate). For practical reasons about ten mass spectra should be measured per peak to preserve peak shape and improve peak deconvolution of overlapping peaks (mathematically pulled apart) [32]. Peaks are deconvoluted based on mathematical algorithms that identify slight, consistent changes in the mass spectra of a convoluted GC peak. Since the ^2D peaks are only a few hundred milliseconds wide, a fast acquisition rate is crucial. It is still possible to use a quadrupole MS in conjunction with GCxGC, but in order to compensate for the slower acquisition rate of this detector, it might be necessary to increase the analysis time (increase peak widths) [27]. A shorter analysis time is only possible (for GCxGC – quadrupole MS) when the range of masses scanned is very small.

TOFMS is based on the relationship between the velocity and mass of an electrostatically accelerated ion [31]. Ionization / fragmentation of molecules result in ions which are accelerated and travel through a flight tube. This flight tube is under a high vacuum (1×10^{-7} Torr) and no further electrostatic field is applied. The time taken for an ion to travel through the flight tube towards the detector is measured. This gives an indication of its mass, since heavier ions will travel slower than lighter ions. Small deviations in the initial starting position of ions in the flight tube and small differences in their individual kinetic energies cannot be avoided. To compensate for this difference in kinetic energy and to improve mass resolution, a reflectron (or ion reflector) is used to align and reflect ions, after which they travel to the detector [31]. The reflectron effectively lengthens the flight path and the flight time is increased by a larger fraction than the flight path.

The detector consists of a metal plate with microchannels, each acting as an electron multiplier, giving sufficient current for the detection of each ion's collision with the plate [31]. The resulting high-speed electrical signal is amplified and

recorded. The mass-to-charge ratio (m/z) can consequently be determined from time-of-flight or arrival time (related to the velocity) of each ion. These m/z ratios are used to compile a mass spectrum for each compound by looking at the distribution of ions. The mass spectrum of a compound consists of the different mass-to-charge ratios of ions (fragments) together with their relative intensities. For each compound this distribution of ions is unique, since each compound will fragment in a unique way. Note that the mass spectra of isomers are often very similar.

When combined with TOF-MS, the identification of compounds and compound classes become much more reliable, since a third dimension of information (mass spectral data) is added to the two-dimensional GCxGC system [28]. The mass spectrum obtained for each compound can be compared to commercial libraries (e.g. from the National Institute of Standards and Technology; NIST) to get an indication of the identity of the compound. This, together with the relative position of the peak on the two-dimensional chromatogram (retention time in 1D and 2D), can be used to tentatively identify compounds. Final confirmation of a compound's identity always requires the injection of the indicated standard.

3.2 *Data analysis*

The 2D separations are stacked next to one another to obtain a three dimensional image as shown in Figure 3.6. When this surface is viewed from above, it can be seen as a two dimensional contour plot and this is usually the way in which GCxGC chromatograms are represented (see Figure 3.7). In the contour plot blue indicates the baseline and red indicates a very high peak intensity.

Software is used to integrate each of these peaks that are above a specified signal-to-noise ratio. For GCxGC-FID, the response signal is used, whilst the total ion signal (TIC) is used for GCxGC-TOFMS. With GCxGC-TOFMS peaks that partially (but not perfectly) overlap can be separately detected or deconvoluted provided that they have different mass spectra. This is achieved with sophisticated computational routines using chemometrics [32].

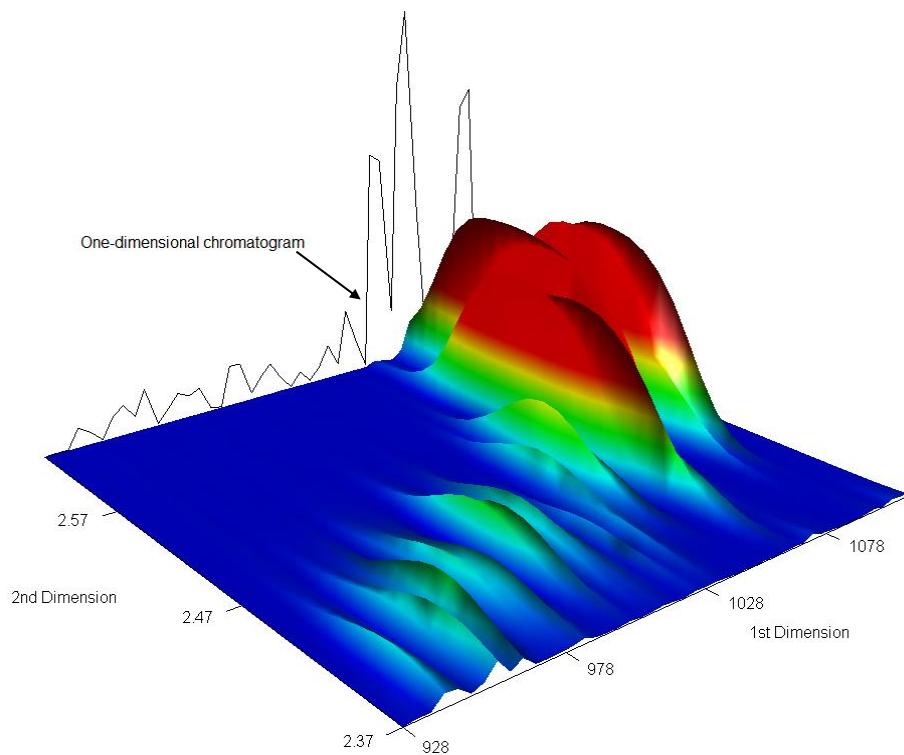


Figure 3.6 The one-dimensional chromatograms, obtained during each modulation period, are placed next to one another to form the two dimensional chromatogram. The third dimension is the intensity of the peaks. The one dimensional chromatogram shown, indicates the separation achieved in the first dimension.

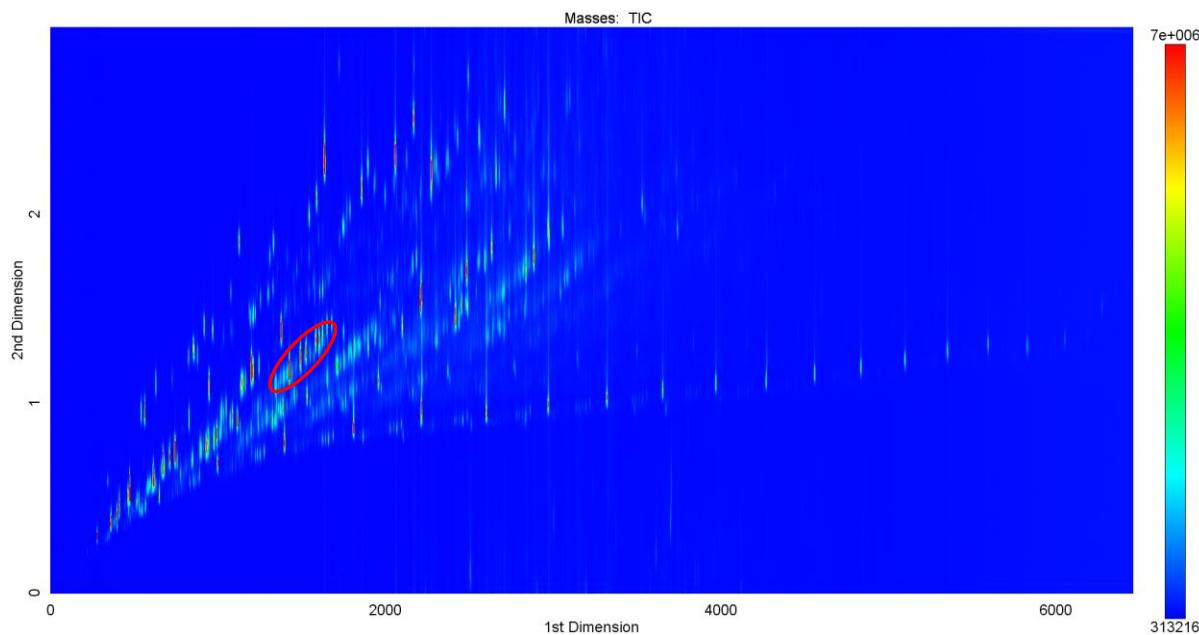


Figure 3.7 An example of a two-dimensional contour plot. This is used to represent the three-dimensional data obtained from GCxGC analysis in two dimensions. Each bright line on the blue background (if expanded it looks like an ellipse) represents a chromatographic peak. Compounds with similar molecular features tend to group together. The red ellipse indicates the position of a “class of compounds”. These classes are ordered like roof tiles in the two-dimensional space.

For two consecutive ^2D separations, the retention time of a single component may vary. When a ^1D peak is modulated for the first time, it is separated in the ^2D at a certain temperature. At the time of the second modulation (of the same ^1D peak), the temperature has increased slightly and thus the same compound will elute a little earlier in the ^2D as illustrated in Figure 3.8. It is possible for the software to identify such peaks and treat them as one peak during integration procedures. Note that this should not be confused with the deconvolution procedures described previously, where compounds with different mass spectra can be distinguished from one another.

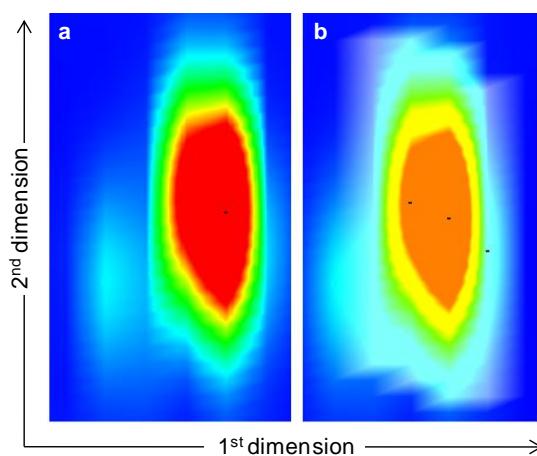


Figure 3.8 Combination of ^2D peaks of a single compound which elute at slightly different times. The same peak is shown in (a) and (b). The black dot is a peak marker and it can be seen that the peak in (a) actually consists of three peaks as shown in (b). The software automatically combines these peaks (from three consecutive modulations) and represents it as one peak (a). The peak eluting from the first dimension, was thus modulated three times in this case.

3.2.1 Classification of compounds into chemical groups

If the separation system is optimised, compounds with similar structure tend to group together in the two dimensional chromatogram (see Figure 3.7), providing class-ordered chromatograms [21]. Thus the qualitative and quantitative composition of the sample can be expressed in terms of groups of compounds, which is sufficient for the aim of this project and much simpler than identifying each component individually [33]. This approach is used often in literature due to the complex nature of petrochemical samples and is considered to be one of the main advantages of GCxGC [21 and references therein].

In literature, GCxGC is often used to distinguish between saturates (n- and iso-paraffins as well as naphthenes), monoaromatics, di-aromatics and tri-aromatics in petrochemical samples [34,35]. These compounds are arranged in readily distinguishable groups on the two-dimensional chromatogram. These larger groups can be classified into smaller groups, especially when combined with TOFMS as detection method [28].

According to Beens and Brinkman [36] petroleum consists of two major types of compounds: hydrocarbons and hetero-compounds. The hydrocarbon group consists of acyclic and cyclic alkanes, alkenes and aromatics. The hetero-compounds can be divided into sulphur, nitrogen, oxygen and metal-containing compounds [36]. A similar approach was described by Schoenmakers et. al. [30]. The polar compounds present in aviation fuels have also been classified into phenols, quinolines, carbazoles, amines, indoles, pyridines and sulphur compounds [37].

4. Properties of two-dimensional separation systems

Two-dimensional separation systems are ideal for the analysis of complex mixtures. This is due to intrinsic characteristics of such systems, including the enhanced peak capacity and orthogonality.

4.1 Enhanced peak capacity

Consider a GCxGC system where the first dimension separates compounds based on volatility and the second dimension separates compounds based on polarity; the fractions that are reinjected onto the second column thus contain several compounds with similar vapour pressures. These compounds are then further separated according to polarity in the second dimension. For this reason the peak capacity of a GCxGC system is defined as the product of the peak capacities of the first and second columns [38]. To exploit this gain in peak capacity fully, it is necessary for the system to be *orthogonal*.

4.2 Orthogonality

A two dimensional system is said to be *orthogonal* when the two dimensions are independent of one another, i.e. the mechanism of separation in the first dimension is sufficiently different from the mechanism of separation in the second dimension [39,40]. This can be illustrated by considering the example where compounds are separated based on volatility in the first dimension. The fractions that are reinjected onto the second column contain several compounds with similar vapor pressures. If the second dimension also separates compounds based on volatility, no additional separation is obtained and thus the compounds will not be further resolved. The peaks will thus elute in a diagonal line across the two-dimensional separation space.

For an orthogonal system, the peaks will be distributed throughout the two-dimensional separation space. The peak locations in this space are defined by the first and second dimension retention times (1t_R and 2t_R respectively). The *orthogonality* of a system can be quantified using the Pearson's correlation coefficient [41,42], which gives an indication of the linear association present (see Chapter 4, Section 5) between 1t_R and 2t_R . If there is a strong linear association (e.g. when peaks are aligned on the diagonal of the separation space), the system is not orthogonal. If the system is orthogonal or near-orthogonal there will be little or no linear association between 1t_R and 2t_R . Other, more complicated methods for measuring *orthogonality* have been proposed in literature [43].

5. Analysis of petrochemical samples

Currently GCxGC (especially when combined with TOFMS) is the only analytical technique capable of giving detailed chemical information regarding the composition of complex mixtures such as petrochemicals [3,21,33,44] and FT products [45,46]. This is due to the enhanced peak capacity and ordered chromatograms that can be obtained. GCxGC is especially useful when working with petrochemical samples in the middle distillate range e.g. jet fuel (kerosene) and diesel, because the boiling point ranges of such samples coincide with the working conditions of the GCxGC

instrument. It has been reported that GCxGC can successfully be used to analyse and quantitate aromatics [47] and oxygenates [48,49] in fuels.

6. Conclusion: the importance of chemical characterisation

The proposed chemical characterisation of petrochemicals by GCxGC is very important since a strong relationship exists between the chemical composition and physical properties of fuels. The physico-chemical relationship of several diesel and jet fuel properties have been investigated such as cetane number, cloud point and specific gravity [50,51]. Models have been proposed for the prediction of these physical properties based on the chemical composition of the fuel in question [50,51]. In order to construct such a model, chemometrics is often used and some of these statistical methods will be discussed in more detail in the following chapter.

References

1. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*; Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, p. 675.
2. De Zeeuw, J.; Luong, J. *Trends in Analytical Chemistry* **2002**, 21, pp. 594 – 607.
3. Phillips. J.B.; Xu, J. *Journal of Chromatography A* **1995**, 703, pp. 327 – 334.
4. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, pp. 51 – 52.
5. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*; Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, p. 688 – 689.
6. Martin, A.J.P.; Synge, R.L.M. *The Biochemical Journal* **1941**, 35, p. 1358 – 1368.
7. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*; Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, p. 681 – 687.
8. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, pp. 75 – 86.
9. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, pp. 48 – 49.
10. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*; Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, p. 679 – 680.
11. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, p. 62.
12. Peters, D.G.; Hayes, J.M.; Hieftje, G.M. *Chemical separations and measurements: The theory and practice of analytical chemistry*; W.B. Saunders Company: United States of America, **1974**, pp. 568 – 571.
13. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*; Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, p. 691.

14. Blumberg, L.M.; Klee, M.S. *Journal of Chromatography A* **2001**, 918, pp. 113 – 120.
15. Skoog, D.A.; Holler, F.J.; Nieman, T.A. *Principles of instrumental analysis – 5th Edition*, Brooks/Cole Thomson Learning Inc.: United States of America, **1998**, pp. 692 – 693.
16. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, pp. 97 – 99.
17. Peters, D.G.; Hayes, J.M.; Hietje, G.M. *Chemical separations and measurements: The theory and practice of analytical chemistry*; W.B. Saunders Company: United States of America, **1974**, pp. 536 – 541.
18. Miller, J.M. *Chromatography concepts and contrasts – 2nd Edition*; John Wiley & Sons Inc.: United States of America, **2005**, pp. 94 – 96.
19. Giddings, J.C. *Analytical Chemistry* **1967**, 39, pp. 1027 – 1028.
20. Giddings, J.C. *Journal of Chromatography A* **1995**, 703, pp. 3 – 15.
21. Dalluge, J.; Beens, J.; Brinkman, U.A.Th. *Journal of Chromatography A* **2003**, 1000, pp. 69 – 108.
22. Schoenmakers, P.; Marriott, P.; Beens, J. *LC GC Europe* **2003**, 16, pp. 335 – 339.
23. Murphy, R.E.; Schure, M.R.; Foley, J.P. *Analytical Chemistry* **1998**, 70, pp. 1585 – 1594.
24. Phillips, J.B.; et al. *Journal of High Resolution Chromatography* **1999**, 22, pp. 3 – 10.
25. Marriott, P.J.; Kinghorn, R.M. *Trends in Analytical Chemistry* **1999**, 18, pp. 114 – 125.
26. Pursch, M.; Eckerle, P.; Biel, J.; Streck, R.; Cortes, H.; Sun, K.; Winniford, B. *Journal of Chromatography A* **2003**, 1019, pp. 43 – 51.
27. Frysinger, G.S.; Gaines, R.B. *Journal of High Resolution Chromatography* **1999**, 22, pp. 251 – 255.
28. Van Deursen, M.; Beens, J.; Reijenga, J.; Lipman, P.; Cramers, C.; Blomberg, J. *Journal of High Resolution Chromatography* **2000**, 23, pp. 507 – 510.
29. Blomberg, J.; Riemersma, T.; Van Zuijlen, M.; Chaabani, H. *Journal of Chromatography A* **2004**, 1050, pp. 77 – 84.
30. Schoenmakers, P.J.; Oomen, J.L.M.M.; Blomberg, J.; Genuit, W.; Van Velzen, G. *Journal of Chromatography A* **2000**, 892, pp. 29 – 46.

31. De Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications – 3rd Edition*; John Wiley & Sons Ltd.: Great Britain, **2007**, pp. 126 – 133.
32. Pierce, K.M.; Hoggard, J.C.; Mohler, R.E.; Synovec, R.E. *Journal of Chromatography A* **2008**, 1184, pp. 341 – 352.
33. Blomberg, J.; Schoenmakers, P.J.; Beens, J.; Tijssen, R. *Journal of High Resolution Chromatography* **1997**, 20, pp. 539 – 544.
34. Beens, J.; Boelens, H.; Tijssen, R.; Blomberg, J. *Journal of High Resolution Chromatography* **1998**, 21, pp. 47 – 54.
35. Vendeuvre, C.; Ruiz-Gerrero, R.; Bertoncini, F.; Duval, L.; Thiébaut, D. *Oil & Gas Science and Technology – Revue d’ IFP* **2007**, 62, pp. 43 – 55.
36. Beens, J.; Brinkman, U.A.Th. *Trends in Analytical Chemistry* **2000**, 19, pp. 260 – 275.
37. Striebich, R.C.; Contreras, J.; Balster, L.M.; West, Z.; Shafer, L.M.; Zabarnick, S. *Energy & Fuels* **2009**, 23, pp. 5474 – 5482.
38. Giddings, J.C. *Analytical Chemistry* **1984**, 56, pp. 1258A – 1270A.
39. Venkatramani, C.J.; Xu, J.; Phillips, J.B. *Analytical Chemistry* **1996**, 68, pp. 1486 – 1492.
40. Ryan, D.; Morrison, P.; Marriott, P. *Journal of Chromatography A* **2005**, 1071, pp. 47 – 53.
41. Liu, Z.; Patterson, D.G.; Lee, M.L. *Analytical Chemistry* **1995**, 67, pp. 3840 – 3845.
42. Zhu, S. *Journal of Chromatography A* **2009**, 1216, pp. 3312 – 3317.
43. Van Gyseghem, E.; et al. *Journal of Pharmaceutical and Biomedical Analysis* **2006**, 41, pp. 141 – 151.
44. Von Mühlen, C.; Zini, C.A.; Caramão, E.B.; Marriott, P.J. *Journal of Chromatography A* **2006**, 1105, pp. 39 – 50.
45. Bertoncini, F.; Marion, M.C.; Brodusch, N.; Esnault, S. *Oil & Gas Science and Technology – Revue d’ IFP* **2009**, 64, pp. 79 – 90.
46. Vendeuvre, C.; Bertoncini, F.; Duval, L.; Duplan, J-L.; Thiébaut, D.; Hennion, M-C. *Journal of Chromatography A* **2004**, 1056, pp. 155 – 162.
47. Frysinger, G.S.; Gaines, R.B.; Ledford Jr, E.B. *Journal of High Resolution Chromatography* **1999**, 22, pp. 195 – 200.

48. Adam, F.; Bertoncini, F.; Coupard, V.; Charon, N.; Thiébaut, D.; Espinat, D.; Hennion, M-C. *Journal of Chromatography A* **2008**, 1186, pp. 236 – 244.
49. Frysinger, G.S.; Gaines, R.B. *Journal of High Resolution Chromatography* **2000**, 23, pp. 197 – 201.
50. Cookson, D.J.; Smith, B.E. *Energy & Fuels* **1990**, 4, pp. 152 – 156.
51. Cookson, D.J.; Smith, B.E. *Energy & Fuels* **1992**, 6, pp. 581 – 585.

Chapter 4: Chemometrics

Contents

1. Chemometrics.....	47
2. What is a <i>model</i> ?	47
3. An overview of basic statistics	47
4. Linear Regression.....	50
5. Linear Association	53
5.1 Association vs. Causality	55
6. Working with mixtures.....	55
6.1 Representing composition data	55
6.2 Mixture models	56
References	59

1. Chemometrics

“An experiment is any action or process whose outcome is subject to uncertainty” [1]. This statement is especially true for a chemical experiment. In chemistry the one certainty is that a significant amount of data will likely be generated. Wold [2] describes chemometrics as “the art of extracting chemically relevant information from data produced in chemical experiments”. *Models* are often used to evaluate and represent the information hidden in chemical data, and thus a basic knowledge of statistics is necessary. Since a chemical system is normally very complex, the data obtained from chemical experiments is often multivariate. Thus multivariate statistics play a crucial role in chemometrics [2]. Chemometrics is especially useful for handling and chemically interpreting the large amounts of data generated with comprehensive two-dimensional separations [3] for the sake of finding correlations between the chemical composition of a mixture and its physical properties.

2. What is a *model*?

Wold [2] describes a *model* (M) as a way of defining the relationship between *experimental values* (y). A *model* always has a certain degree of *variability / noise* (ε) associated with it. Thus: “ $y = M + \varepsilon$ ” or “Data = Chemical Model + Noise” [2]. The *chemical model* will be referred to as the *model*, whilst the *variability / noise* will be referred to as the *error / residual* (in the *model*).

3. An overview of basic statistics

Statistics can be described as “the art of drawing conclusions from data, and making decisions, in the presence of variability” [2]. The terminology used in statistics is explained in Table 4.1.

In chemistry the idea of a *process* is more readily understandable than that of a *population* [4]. Usually a *sample* is taken from a *population / process* and its *statistics* is determined (see Table 4.2). Once the *statistics* of a *sample* is known, it can be used to estimate the *parameters* of the *population / process*. This is called statistical inference [5].

Table 4.1 Basic terminology used in statistics [4] and examples of how it can be applied in chemistry.

Statistical term	Definition	Chemistry example
Population	“All possible items or units possessing one or more common characteristics under specified experimental or observational conditions”	The refining process
Process	“A repeatable series of actions that results in an observable characteristic or measurement”	
Variable	“A property or characteristic on which information is obtained in an experiment”	Lubricity or composition of a diesel fuel produced by a refining process The number of variables is denoted by “ k ”
Observation / observed value	“The collection of information in an experiment, or actual values obtained on variables in an experiment”	Chemical analysis of a diesel stream and the results obtained from this analysis. HFRR analysis for the measurement of lubricity and the WSD values obtained. The number of observations is denoted by “ n ”
Response variable	“Any outcome or result of an experiment”	Relative area percentage from GCxGC results
Factors	“Controllable experimental variables that can influence the observed values of response variables”	Type of catalysts used in refining processes. Conditions of chromatographic separation, e.g. identity of stationary phase
Sample	“A group of observations taken from a population or a process”	A group of specified diesel streams
Parameters	“A numerical characteristic of a population or a process”	Average lubricity or average composition of streams produced by a refining process
Statistics	“A numerical characteristic that is computed from a sample of observations”	Average lubricity or average composition of a group of specified diesel streams

Since it is impossible to analyse all the possible streams produced by refining processes, a *sample* of streams (a set of streams obtained from a refinery) can be analysed and the *sample statistics* from this data can be used to estimate the *parameters* of the process. *Summary statistics* are often used to give an overview of the *variability* in the *sample* [1,5]. The most commonly used *summary statistics* are revised in Table 4.2.

Table 4.2 Summary statistics that can be used to give an overview of the variability in the sample [1,5].

Summary statistic	Mathematical expression and description	Parameter to be estimated
Mean	$\bar{x} = \frac{\sum x_i}{n}$ <p>The central value for a normally distributed data set. It can be used to estimate the mean of the <i>population</i> (μ).</p>	μ
Median	The central value, i.e. 50% of all the values are below the median and 50% of all the values are above it. It is not as easily influenced by outliers as the mean.	
Mode	The value that appears most frequently in the data set.	
Standard Deviation	$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$ <p>The spread of results. It estimates the probable error of a single measurement. Describes how far the values are from the mean.</p>	σ
Variance	$s^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$ <p>The spread of results. It estimates the probable error of a single measurement. Describes how far the values are from the mean.</p>	σ^2
Standard deviation of the mean / Standard error	$s_{mean} = \frac{s}{\sqrt{n}}$ <p>Describes the precision of the mean. It estimates the probable error of the mean</p>	
Range	The difference between the largest and smallest <i>observed values</i> . The range can be used to describe the spread of results for small data sets.	
Degrees of freedom	The number of values in the final calculation of a statistic that are free to vary.	

Probability theory is the mathematical foundation of statistics [6]. When doing statistical analyses of data, a hypothesis (*null hypothesis*, H_0) is often proposed and the probability of the hypothesis being true is tested. Because of the variability that is always associated with a data set, it is necessary to quantify the certainty with which conclusions can be made from the data. This is called the *confidence level* ($1 - \alpha$), and α is referred to as the *significance level*. If the *null hypothesis* is rejected (with a certain level of *confidence*), the *alternative hypothesis* (H_A) is accepted with probability $1 - \alpha$. An example of a *null hypothesis* is that the *population mean* (μ) is a specified value, thus the *alternative hypothesis* is that the *population mean* is not equal to the specified value [6]. This can be written as follows:

$$H_0: \mu = 20$$
$$H_A: \mu \neq 20 \text{ or } H_A: \mu > 20 \text{ or } H_A: \mu < 20$$

The *sample mean* (\bar{x}) is calculated using the formula in Table 4.2. The *sample mean* is used to estimate the *population mean* (μ). If the *null hypothesis* is rejected at a *significance level* of 5 % the *alternative hypothesis* is accepted with 95 % confidence. As seen above, the *alternative hypothesis* can be two-sided ($\mu \neq 20$) or one-sided ($\mu > 20$ or $\mu < 20$).

There are two types of errors that need to be considered when testing a hypothesis; Type I and Type II errors [6]. A Type I error is a false positive result, i.e. decision to reject H_0 even though it is true. A Type II error is a false negative result, i.e. decision to accept H_0 even though it is false.

4. Linear Regression

Often, the variation in one *variable* can be described by the variation in another *variable*, e.g. boiling point (the *dependent variable*, y) increases with an increase in molecular weight (the *independent variable*, x). Linear regression is the process of determining the best fitting straight line for such a data set [5-7]. The equation of this straight line (Equation 4.1) is called the *model* that explains the variation in the data.

To formulate a *model* the values of β_0 (the intercept) and β_1 (the slope) in the straight line equation are estimated. The third term (ε) represents the *error* in the *model* (see Chapter 4, Section 2).

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (4.1)$$

To determine the values of β_0 and β_1 , several methods have been developed, including the least squares method, minimum variance method and the maximum likelihood method [7].

The straight line regression equation is given below (Equation 4.2) and illustrated in Figure 4.1.

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\text{SST} = \text{SSM} + \text{SSE} \quad (4.2)$$

Where SST is the sum of squares of total *variability* in the data, SSM is the sum of squares attributable to the *model* and SSE is the sum of squares of *errors*. Note that y_i is the observed value (where $i = 1, 2, \dots, n$), \hat{y}_i is the predicted value (using Equation 4.1) and " $y_i - \hat{y}_i$ " is the *residual / error* (ε). By using the sum of squares, the influence of negative values is accounted for. The SSE represents the *variability* in the *model* (sum of squares of ε). In the least squares approach this term is minimised, i.e. β_0 and β_1 is determined in such a way that the differences between the actual and predicted values (SSE) are as small as possible. The SSE is represented in Figure 4.1 by the sum of the areas of the blue squares.

There are two ways to test the validity of the *model*; a *null hypothesis* that the slope is equal to zero ($H_0: \beta_1 = 0$) is tested, followed by testing the *null hypothesis* that the intercept is equal to zero ($H_0: \beta_0 = 0$) [7]. If the hypothesis that $H_0: \beta_1 = 0$ is rejected it implies a *model* that includes x , is better for predicting y than a *model* that does not include x with some probability. The hypothesis that $H_0: \beta_0 = 0$ is only considered when experience or theory can explain that the straight line should go through the

origin [7]. It is also important to test whether the *residuals* are normally distributed, since this is an indication of whether the *error* in the *model* is random or not, and is a necessary for hypothesis testing [6,7].

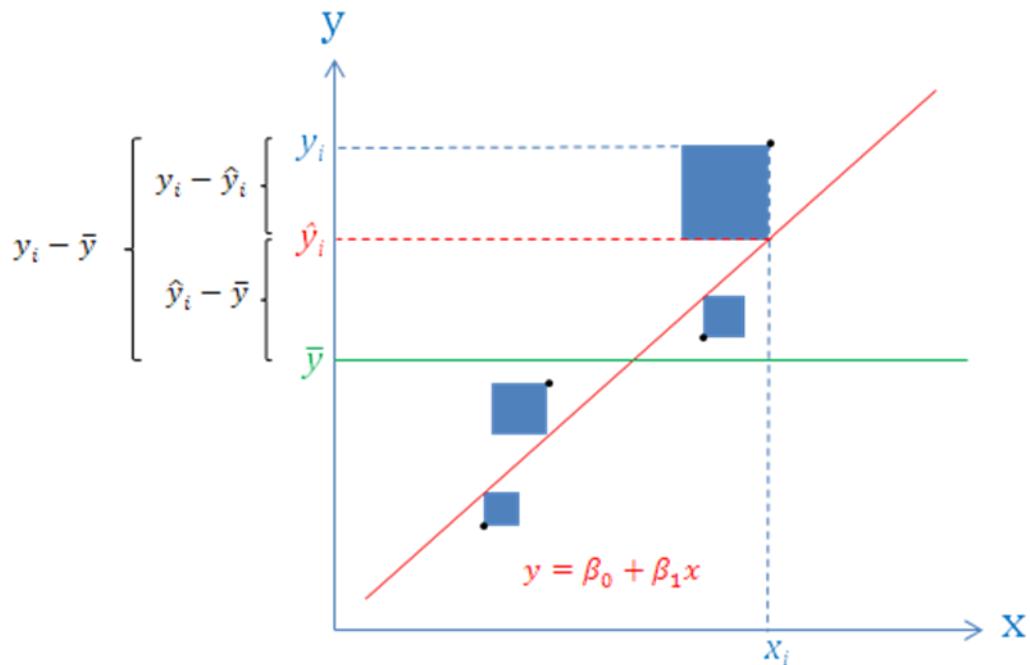


Figure 4.1 A graphical explanation of linear regression [6]. The black dots represent the data points and the red line represents the predicted regression equation. The sum of the areas of the squares represents the *variability in the model* (SSE) which is to be minimised when the least squares approach is used to find the best fitting line through the data points [6].

Linear multiple regression is an extension of the straight-line regression analysis described above. In this case the *dependent variable* (y ; e.g. lubricity) is dependent on more than one *independent variable* (x_1, x_2, \dots, x_k , e.g. composition or concentration of different constituents of a stream). Thus several *parameters* ($\beta_0 + \beta_1 + \dots + \beta_k$) need to be estimated (Equation 4.3), and the relationships between the *independent variables* need to be taken into account [7].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (4.3)$$

It is important to keep in mind that a *model* can only be obtained if the number of *independent variables* (k) are less than or equal to the number of *observations* (n) minus one, i.e. $k \leq n - 1$ [7].

5. Linear Association

Association can be defined as the degree of dependence between two *variables* [5,7]. The main difference between linear association and linear regression is that in linear regression it is assumed that there is no *error* in x , whilst in linear association, *error* is allowed for in x and y .

To better understand this concept, consider a plot of y (e.g. boiling point) vs. x (e.g. molecular weight). The distribution of the data points on the graph gives an indication of the relationship between the *variables*, e.g. if the points follow a linear trend (in any direction) a linear relationship clearly exists between the boiling point and the molecular weight (see Figure 4.2). If the points are encircled, an ellipse is obtained. A narrower ellipse indicates a stronger linear association. If the points are scattered more randomly, the ellipse enclosing these points will become wider, indicating less linear association (or dependence). When a circle is obtained it means that there is no linear association present [6].

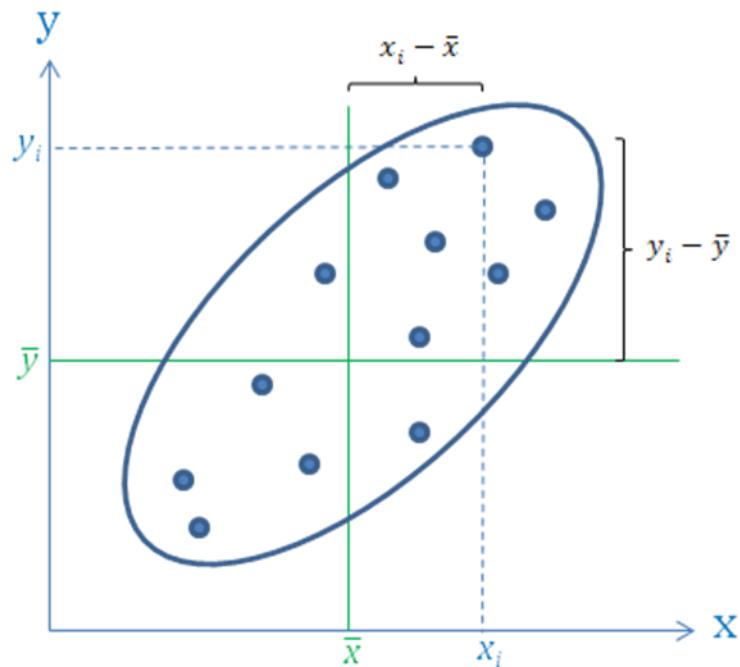


Figure 4.2 A graphical explanation of linear association. Variation in x and y is considered (compare to linear regression where variation is only allowed for in y).

Pearson's correlation coefficient (r) is a measure of the linear association between two variables, x and y . It is always true that $-1 \geq r \geq 1$ and if $r = 0$ there is no linear association between the two *variables*. Larger $|r|$ values indicate a larger association between the *variables*. The sign of the *correlation coefficient* indicates whether a positive or negative association exists [7]. The mathematical equation used to calculate the *correlation coefficient* is:

$$r = \sum \frac{(x_i - \bar{x})(y_i - \bar{y})}{ns_x s_y} \quad (4.4)$$

The *square of the correlation coefficient* (r^2) is a quantitative measure of the improvement in the fit by using x , i.e. can we predict the value of y better when the effect of x is taken into account [7]. The *square of the correlation coefficient* is a more conservative measure of association and thus might be better to assess the strength of the linear association between x and y .

It is important to realize that the magnitude of the slope of the regression line is not measured and the *correlation coefficient* is not an indication of whether or not the straight-line model is appropriate [7].

The question remains whether this *correlation* (or linear association) is statistically significant. Let the *correlation coefficient* determined by equation 4.4 be r_1 . If no linear relationship exists between the variables, let the *correlation coefficient* be r_2 . The probability of obtaining $|r_2| > |r_1|$ by chance alone is determined. If this probability is very small (< 0.05 for a 95% *confidence*) then the *correlation* (or linear association) can be said to be statistically significant [6 and JMP® 8.0.1 Help file].

When multiple variables are present in the data set, a *correlation matrix* containing pairwise *correlation coefficients* can be obtained, i.e. the *correlation coefficients* of all the possible combinations of *variables*.

5.1 Association vs. Causality

An important factor to consider is that even though the *correlation coefficient* may indicate a strong linear association between x and y , this does not necessarily mean that x causes an increase or decrease in y (causality). Experimental proof is needed to conclude that a causal relationship exists. Alternative methods also exist when experimental proof cannot be obtained but these methods are not usually used in scientific studies [7].

6. Working with mixtures

Compositional data has to be treated in a specific way, since the *observed values* all add up to a certain value (usually 1 or 100). The effect of the individual components and their interactions need to be taken into account [8,9]. Specific methods are required to represent such data graphically and to build models to predict properties from the data.

6.1 Representing composition data

The simplest way to represent a composition in two dimensions is by using ternary plots. An example of a ternary plot is given in Figure 4.3 below. The three axes of the plot represent three different components in the mixture (A, B and C). To simplify the discussion, it is assumed that the mixture consists of only three components. For more complex mixtures, the components can often be grouped together into three groups to simplify the interpretation of results.

The position of a point in the triangle represents the mixture's composition. All the possible combinations of A, B and C are limited to the space within the triangle, e.g. it is not possible to have $A = 0.7$, $B = 0.2$ and $C = 0.2$ since these components always have to add up to one. The composition can also be considered in terms of percentage in which case it will add up to 100.

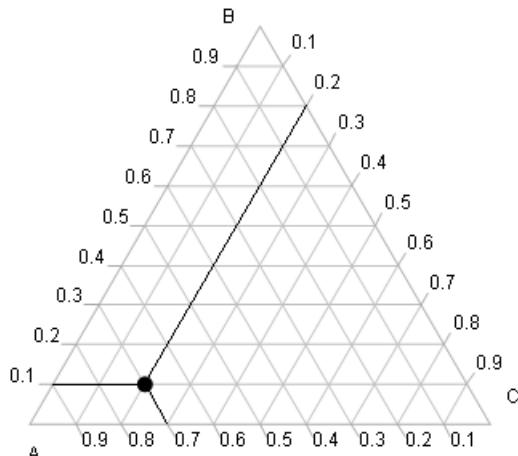


Figure 4.3 An example of a ternary plot and how it can be used to visualize the composition of a mixture. The data point shown indicates a mixture containing 70 % of A, 10 % of B and 20 % of C.

If a point is on the A axis it means the mixtures contains only A and C. Similarly on the B axis, the mixture contains only B and A, whilst on the C axis the mixture contains only C and B. The corners of the triangle represent pure compounds, i.e. there are none of the other compounds present. A data point in the top corner represents 100 % B, in the left corner represents 100 % A and in the right corner represents 100% C.

6.2 Mixture models

Consider a set of nine mixtures as represented by the ternary plot in Figure 4.4. Assume that these mixtures all have a common physical property that can be explained by their chemical composition. Thus every data point on the ternary plot, has a specific response value associated with it that describes the common physical property.

Imagine another axis that is perpendicular to the ternary graph. The response values for each data point can be plotted in this dimension. Since the response values are spread out in a plane and not on a line, a response surface is calculated instead of a response curve (e.g. a straight line or parabola, see Chapter 4, Section 4). This is shown graphically in Figure 4.5 and the response surface can be described mathematically as shown in Equation 4.5 [8,9].

$$y = \beta_1 A + \beta_2 B + \beta_3 C + \beta_4(AB) + \beta_5(AC) + \beta_6(BC) \quad (4.5)$$

Where A, B and C are the components in the mixture and the terms AB, AC and BC represent the interactions among these components [8,9]. Note that there is no intercept in this model. The coefficients ($\beta_1, \beta_2, \dots, \beta_6$) are determined in order to define the model.

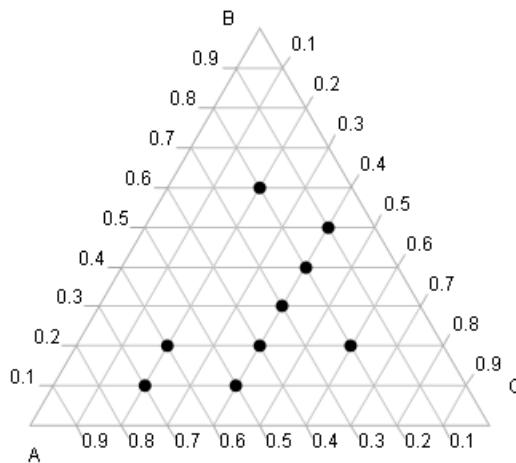


Figure 4.4 The compositions of nine mixtures represented as a ternary plot.

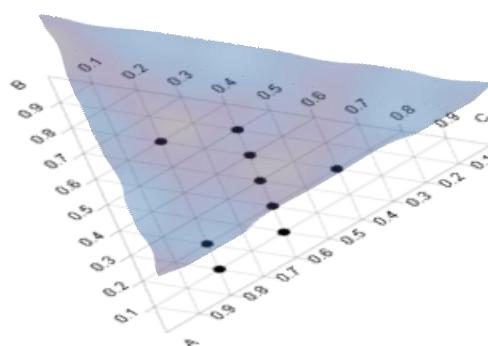


Figure 4.5 The response surface representing a common physical property of the different mixtures.

According to Cornell [9] the coefficients of the linear terms (β_1, β_2 , and β_3) can be interpreted as the fitted response at the corners of the plot, where only one factor (A, B or C) is present, whilst the coefficients on the cross terms (β_4, β_5 , and β_6) indicate the curvature across each edge of the plot (the axes).

Once a model is defined, it can be used to predict the response values (the value of the physical property) of other mixtures by extrapolation or interpolation [9]. It is

always best to use interpolation rather than extrapolation, and therefore the original set of mixtures should have significantly different compositions and pure compounds should also be considered as part of the data set. It is possible to extend above arguments to mixtures containing more than three compounds and this is covered in detail in literature [9]

References

1. Devore, J.L.; Kenneth, N.B. *Modern mathematical statistics with applications*; Thomson Brooks/Cole: California, **2007**, pp. 25 – 37, 50.
2. Wold, S. *Chemometrics and Intelligent Laboratory Systems* **1995**, 30, pp. 109 – 115.
3. Sinha, A.E.; Prazen, B.J.; Synovec, R.E. *Analytical and Bioanalytical Chemistry* **2004**, 378, pp. 1948 – 1951.
4. Mason, R.L.; Gunst, R.F.; Hess, J.L. *Statistical Design and Analysis of Experiments - With Applications to Engineering and Science – 2nd Edition*; John Wiley & Sons: New York, **2003** pp. 9 – 19. Online version available at:
http://knovel.com/web/portal/browse/display?EXT_KNOVEL_DISPLAY_bookid=1406&VerticalID=0
5. Christian, G.D. *Analytical Chemistry – 6th Edition*; John Wiley & Sons: United States of America, **2004**, pp. 74 – 107.
6. Sall, J.; Creighton, L.; Lehman, A. *JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP® – 4th Edition*; SAS Institute Inc.: North Carolina, **2007**.
7. Kleinbaum, D.G.; Kupper, L.L.; Muller, K.E. *Applied regression analysis and other multivariable methods*; Thomson Brooks/Cole Pub.: California, **2008**
8. Scheffé, H. *Journal of the Royal Statistical Society, Series B* **1958**, 20, pp 344 – 360.
9. Cornell, J.A. *Experiments with mixtures: designs, models and the analysis of mixture data – 3rd Edition*; Wiley-Interscience Publication by John Wiley & Sons: United States of America, **1941**, pp. 1 – 12.

Chapter 5: Experimental Setup

Contents

1. Samples	61
2. Lubricity Measurement	61
3. Optimisation of GCxGC system.....	62
3.1 Practical considerations.....	63
3.1.1 Peak shape	63
3.1.2 Temperature limits	64
3.1.3 Detector saturation.....	64
3.1.4 Sample preparation and injection.....	64
3.2 Operational parameters.....	65
3.3 Procedure	67
4. GCxGC-TOFMS	68
4.1 Instrument description	68
4.2 Column selection	68
4.3 Separation conditions	71
4.4 Analysis of samples.....	73
4.4.1 Optimisation of separation on SLB-IL100.....	74
5. Statistical Analyses.....	85
References.....	87

1. Samples

Streams from different refining processes, usually blended or combined in different ratios to give a final diesel product, were investigated. None of these streams contained lubricity additives. Some of the feed streams to the final refining processes were also studied. These streams are summarised in Table 5.1 below. For proprietary reasons, further detail regarding the streams cannot be discussed here.

2. Lubricity Measurement

The HFRR wear scars of the samples were measured at the Fuels Research laboratories in Sasolburg using the HFRR method described previously (see Chapter 2, Section 3.5). The results obtained are shown in Table 5.1. In cases where more than one measurement was taken, the average of these values was used in further calculations. The WSD values show that the feed streams are significantly better lubricants than the final blending streams. This appears to indicate that some compounds that contribute to lubricity are lost during the final refining processes. This reduction in lubricity will be discussed in more detail in later chapters.

Table 5.1 Diesel samples from different refining streams. The lubricity of the samples was measured using an HFRR instrument and measurements are given in terms of WSD.

Sample	Description	WSD (μm)
I	Diesel blending stream	769
II	Diesel blending stream	785
III	Diesel blending stream	525
IV	Diesel blending stream	530
V	Unrefined feed stream to III	244
VI	Unrefined feed stream to IV	235
VII	Diesel blending stream	752
VIII	Diesel blending stream	782
IX	Diesel blending stream	524

3. Optimisation of GCxGC system

Based on the objective of this research a compromise was made during optimisation between class resolution (of especially polar species) and analysis time. The aim of this work was not to find a method which can be automated for the characterisation of diesel samples, but instead to separate and identify compounds or classes of compounds which may contribute to lubricity. Since it was suspected that polar compounds could improve lubricity (see Chapter 2, Section 3.2 and 3.3), it was essential that the polar fractions of the samples were resolved in terms of classes of compounds.

The composition of samples in terms of mass percentage was not determined (for this FID is required together with the use of internal standards), since it was sufficient to use the total ion response as a measure of relative composition. Consequently the results are not directly comparable with GCxGC-FID results. In cases where classes of compounds overlapped, individual peaks were assigned to classes based on mass spectral data (this aspect will be addressed later).

The relative area percentage of each peak was used to give an indication of the amount of the particular compound present. These values were normalised in such a way that the sum of the area of all the sample peaks were 100 %. This method does present some disadvantages but the comparison of relative compositions are supported by the fact that all samples were treated in the same way where data processing and classifications were concerned. It is important to keep in mind that the different columns investigated had varying temperature limits and consequently some of the less volatile compounds might not elute from certain columns. As a result of this, the relative amount present of the more volatile components may be overestimated when using the relative area percentage as an indication of composition.

Some practical aspects, which could also be useful for future work, are explained below. This is followed by a discussion of the three main criteria for an optimised separation within a certain time limit: maximum distribution of compounds throughout

the two-dimensional separation space, maximum class resolution and minimum wrap-around.

3.1 Practical considerations

During the optimisation of the GCxGC system there were several practical aspects that had to be considered. These included peak skewing, the temperature limits of the capillary columns, detector saturation as well as some aspects of sample preparation and injection.

3.1.1 Peak shape

From previous discussions (see Chapter 3, Section 2.4) it was important to acknowledge that peak skewing could occur due to non-linear behaviour at high analyte concentrations and also due to interactions between (lower concentration) polar analytes and exposed silanol groups from the column wall.

Other factors not related to the capillary column could also cause peak-skewing. These factors are known as extra-column effects [1] and are often observed when a “cold spot” is present in the system. Cold spots originate e.g. where there is contact between the column and the sides of the oven. Even though the oven temperature is regulated, the metal on the sides of the oven may take some time to reach the desired temperature. Compounds can also get “trapped” in the inlet or transfer line (to the detector) if these areas are not warm enough, resulting in tailing peaks. To prevent this, these areas should be kept at a higher temperature than the final temperature of the temperature program whenever possible.

Peak tailing or fronting can be observed in both dimensions of GCxGC. The peak shape might sometimes be useful in identifying certain compounds. Consider the situation where the mass spectrum of a compound and its relative position on the two-dimensional chromatogram indicates that it could be a specific aldehyde or analogous ketone. If the peak tails in the ^1D but not in the ^2D where the stationary phases are non-polar and polar respectively, it can be assumed that it is more likely

to be an aldehyde than a ketone, since an aldehyde will more readily form strong hydrogen-bonds with exposed hydroxyl groups from the column wall. If there is no peak tailing in either direction, the compound might be a ketone. These considerations can improve tentative identifications – standards still need to be analysed to identify compounds with absolute certainty.

3.1.2 Temperature limits

The temperature limits of the first and second columns play a significant role in the type of samples that can be analysed. It is important to keep in mind that the second oven is housed inside the first oven and is normally kept at a higher temperature. Thus it is preferred for the second column to have a higher temperature limit than the first, in cases where the first oven temperature needs to be operated at its maximum design temperature.

3.1.3 Detector saturation

The detector can be saturated when a large amount of ions reach the detector at a specific instance. When all the microchannels are discharged, other ions present cannot be detected. This results in relatively lower signals at the peak maximum and thus peak profile distortion (flatter topped peaks). Thus detector saturation should be avoided (by e.g. dilution) at all costs when quantitative analyses are carried out.

3.1.4 Sample preparation and injection

Analytes introduced into a GC system via liquid injection are usually diluted to prevent saturation of the column stationary phase or detector. Complex mixtures such as petrochemical samples present a further challenge in this regard, since many common solvents used for dilution (e.g. hexane) are already present in the sample. Thus it is not feasible to dilute such samples and alternatively a smaller amount of sample can be injected into the column. This is often accomplished by using a *split* injection, where the amount of sample injected into the inlet is split according to a specified ratio and thus only a fraction of the sample enters the GC column.

During trace analysis, a *splitless* injection is preferred over a *split* injection. This is due to the fact that some target compounds may be lost during a *split* injection if proper mixing of compounds in the inlet is not obtained. To establish proper mixing in the inlet, the inlet liner usually contains deactivated glass wool. The glass wool promotes complete evaporation of aerosol droplets, resulting in a non-discriminatory split. Polar compounds can adsorb onto the glass wool (which contains silanol groups) and thus small amounts of these compounds might be lost. Alkanes and less polar compounds do not tend to interact with these silanol groups.

3.2 Operational parameters

For a two-dimensional system to be optimised, three conditions need to be met. Firstly the peaks need to be distributed throughout as much of the two-dimensional space as possible to take advantage of the enhanced peak capacity of the two-dimensional system. This distribution could be used as a measure of *orthogonality* (see Chapter 3, Section 4.2).

Secondly, *classes of compounds* (“roof tiles” in Chapter 3, Section 3.2.1) need to be resolved from one another. Here it is important to distinguish between *compound* resolution and *class* resolution. If compounds that are normally grouped together in a *class* co-elute, it does not influence the overall analysis. *Classes of compounds* that are not sufficiently separated (i.e. “roof tiles” that overlap with one another) may be problematic.

One of the main advantages of GCxGC is that a complex mixture can be characterised in terms of *classes of compounds* present, instead of the individual components (see Chapter 3, Section 3.2.1). When these *classes* are not resolved from one another, identification of individual compounds becomes necessary in order to assign compounds to a *compound class*. This process is simplified when a mass selective detector is used, since extracted ion chromatograms (EIC) can be used to distinguish between *classes of compounds*. This is still a time consuming process and can become impractical when a mass spectrometer as detector is not available.

The last important aspect of a two-dimensional separation is wrap-around which occurs when peaks elute after the modulation period during which it was injected onto the second column. Wrap-around should be avoided or minimised in order to get an optimum separation [2-4]. Wrap-around can only be tolerated when the peaks that are wrapped around elute during the hold-up time in the second dimension (2t_M) or when these peaks do not co-elute with other peaks. Peaks that are wrapped around can be distinguished from those that are not wrapped around, since they have larger peak widths.

There are several ways to prevent wrap-around [2-4]. The simplest way to govern the distribution of peaks in the second dimension is by changing the difference in temperature between the two ovens (see Chapter 3, Figure 3.4, nr 3 and 7). The larger the temperature difference, the higher the temperature of the second oven when a peak elutes from the first column, and thus the quicker it will move through the second column once injected. The rate at which both ovens are heated can also be increased to eliminate wrap-around based on the same principle, but then the first dimension peaks can become too narrow to allow three modulations per peak (see Chapter 3, Section 3.1). More drastic measures, which would have a similar effect, are shortening the second column length or increasing the carrier gas flow rate. When doing this, care should be taken that efficiency (and consequently resolution) is not lost (see Chapter 3, Section 2.1, Equation 3.7).

To prevent wrap-around, the modulation period (allowed time for second dimension separation) can be made longer. Since the second dimension analysis in GCxGC can be considered to be isothermal, the hold-up time in the 2D (the retention time of an unretained compound, 2t_M) can be used to estimate how much time is needed for sufficient separation. A k -value of 3 to 4 is ideal (see Chapter 3, Section 2.1), and thus sufficient resolution should be obtained with a modulation period of four or five times 2t_M (see Equation 3.7). The modulation period should be changed cautiously, since a first dimension peak should still be modulated at least three or four times, to prevent loss of the separation achieved in the 1D . If necessary, the first dimension separation can be slowed down (without loss of resolution) by using a slower temperature programming rate, a process that can lead to extremely long analysis times.

3.3 Procedure

To optimise the two-dimensional separation, mixture separation was first optimised in a one-dimensional system. This was conveniently done by setting up the two columns in series and disabling the modulation of peaks. This resulted in a one-dimensional GC separation based on the first column, with negligible contribution of the second. Once this separation was optimised, the modulation was enabled and consequently a two-dimensional separation was obtained. There are several parameters that can be tuned to optimise this separation, of which the temperature difference between the two dimensions and the modulation period are the most important.

As discussed previously, a group-type analysis of complex mixtures is made possible with GCxGC since structured chromatograms are obtained (see Chapter 3, Section 3.2.1). The chemical groups (to which the chromatographic peaks were allocated) were defined as categories, subcategories, classes and subclasses. A category consisted of a group of compounds with the same functional groups, e.g. aliphatic hydrocarbons (AIHC), aromatic hydrocarbons (ArHC) and oxygenates (Ox). This is similar to what has been done in literature (see Chapter 3, Section 3.2.1 and references therein).

The categories were further divided into subcategories e.g. AIHC was divided into groups with different degrees of saturation such as alkanes, alkenes etc. whilst ArHC was grouped according to the number of aromatic rings present, similar to what has been done in literature [5,6]. Subcategories were divided into classes based on the number of carbons present, e.g. all possible isomers of a C₆-alkane was grouped together in one class and all C₇-alkanes was grouped together in a different class. Thus all the compounds in a certain class have the same molecular formula and are constitutional isomers. Classes were further divided into subclasses, e.g. normal C₆-alkanes and branched C₆-alkanes. Finally the classes or subclasses could be considered in more detail to gain information regarding individual compounds.

4. GCxGC-TOFMS

4.1 *Instrument description*

Chemical analysis was carried out using a LECO Pegasus® 4D GCxGC-TOFMS system (Figure 5.1). The system consists of a 7890A GC from Agilent Technologies, modified to contain a modulator and secondary oven (see Chapter 3, Section 3.1 for information regarding the theory of operation). The modulator can also accommodate a second parallel column set which could be connected to the FID. The GC was fitted with a 7683B series autosampler from Agilent Technologies. The two columns were connected with a presstight connector from Restek®. ChromaTOF® software (version 4.24 and later 4.32) was used to operate the instrument and for data capturing and processing.



Figure 5.1 The LECO Pegasus® 4D GCxGC-TOFMS system [7] used for the chemical analysis of diesel samples.

4.2 *Column selection*

There are two related factors to consider when selecting columns for a two-dimensional system. Firstly, *orthogonality* (see Chapter 3, Section 4.2) is necessary

in order to exploit the enhanced peak capacity of GCxGC. Secondly, the primary objective of the separation will determine in which dimension the selected columns will be placed. This is due to the fact that the first column contains a larger number of theoretical plates (it is much longer) than the second column. Thus the ¹D separation is more *efficient* (see Equation 3.7) than the ²D separation. If, for example, the distinction between different functional groups is very important, it is advantageous to select a column with a stationary phase that has a high selectivity towards functionality for the first dimension (resolution being a function of both selectivity and efficiency, see Equation 3.7). A column with a more polar stationary phase (e.g. PEG or IL) will have a high selectivity towards more polar compounds and can thus be used to distinguish between compounds with different functional groups.

In literature the ¹D column often has a less polar stationary phase and the ²D column has a more polar stationary phase [3]. In this study a “reverse-type” column combination was also considered, where the first column is more polar than the second. This specific column combination has hardly been used in literature, but better class separation is observed as explained above [8,9].

Literature suggests that a column with a stationary phase with a similar chemical nature to that of the analytes be used to prevent the onset of non-linearity at high concentrations and thus peak distortion [10]. For mixtures containing a wide variety of compounds this is not always possible and it might be necessary to compromise, e.g. a longer analysis time may be needed in order to obtain sufficient separation under linear conditions. In some cases peak tailing can be tolerated when the tailing analytes are present in sufficient quantities to produce peaks that can be distinguished from the noise levels. Peak tailing can be problematic during quantitative studies and the use of an internal standard is suggested in such instances. Based on theory, it is vital for the two dimensions of GCxGC to separate compounds based on different mechanisms and thus peak tailing cannot necessarily be avoided in both dimensions. Note that peak fronting will occur more readily in the ²D column of a GCxGC system due to column saturation (a very short, narrow column with a thin film thickness is used – see Chapter 3, Section 2.4 and 3.1).

Taking into account above statements and based on literature cited, four different column sets were used to efficiently separate the wide range of samples (see Table 5.2). Column set 3 can be called a “normal-phase” column combination since the second dimension is more polar than the first (see Chapter 3, Table 3.1). Column sets 1, 2 and 4 have “reversed” polarities. Figure 5.2 illustrates the chemical nature of the stationary phases of the different columns used. The polarity of the different stationary phases increase in the order: a << b < c < d. Columns ZB-5 and ZB-5MSi can be considered to be essentially non-polar, whilst the other columns have more important polar characteristics as explained in Chapter 3, Section 2.3.

Table 5.2 Summary of the four different column sets used to separate the diesel samples.

Column description		Column set			
		1	2	3	4
First Dimension (¹ D)	Phase	RXi-17	ZB-Waxplus	ZB-5	SLB IL100
	Length (m)	30	60	30	30
	¹ d _c (μm)	250	250	250	250
	¹ d _f (μm)	0.25	0.25	0.25	0.20
Second Dimension (² D)	Phase	ZB-5MSi	ZB-5MSi	RXi-17	ZB-5MSi
	Length (m)	2	2	1	2
	² d _c (μm)	100	100	100	100
	² d _f (μm)	0.1	0.1	0.1	0.1

The SLB-IL100 column has only recently been commercialised. It has been proven to be very useful in biodiesel analysis [11] and thus was considered in this study as well. Peak tailing is expected for very polar compounds (e.g. alcohols, carboxylic acids and amines) due to the nature of the anion [12]. The major advantage of using an ionic liquid as a stationary phase is its high polarity and high thermal stability [12]. It is possible to make stationary phases for capillary columns with ionic liquids consisting of various cations and anions [12]. By doing this the selectivity of the column can be tuned to specific needs.

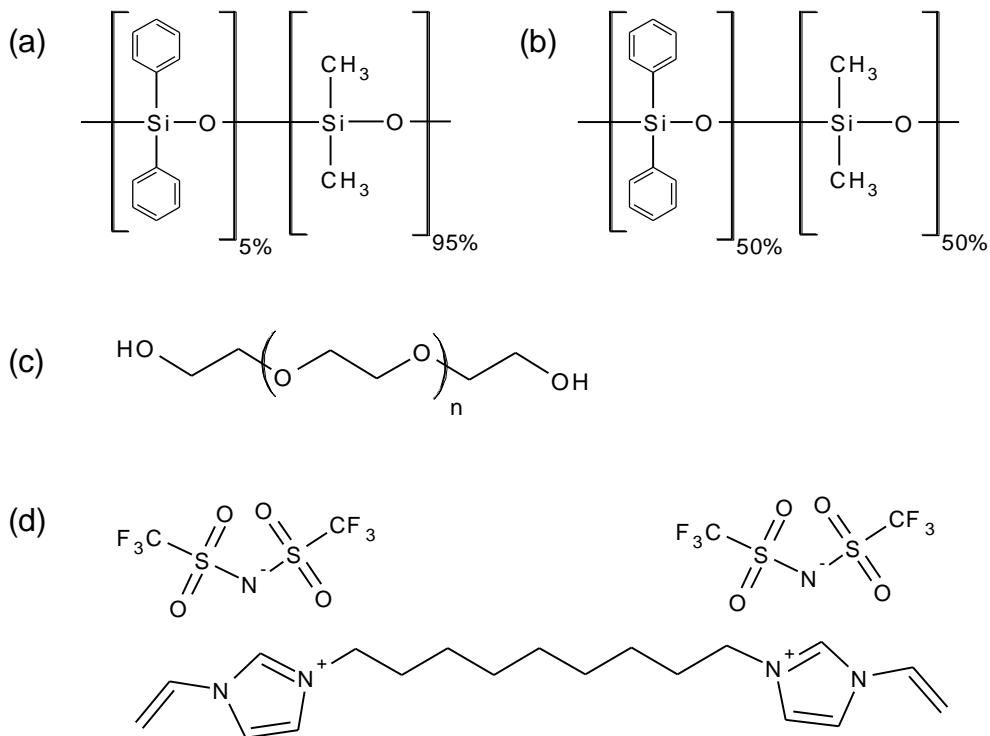


Figure 5.2 The chemical composition of the stationary phase of (a) ZB-5 (ZB-5MSi is very similar), (b) RXi-17, (c) ZB-Waxplus and (d) SLB-IL100.

4.3 Separation conditions

To optimise sample separation, several GC parameters could be changed as required. Due to the varying composition of the samples, different methods (sets of parameters) were used to analyse the different samples. Some of the GC and MS parameters are discussed below and the temperature programming and modulator settings are summarised in Table 5.3. After sample analysis, data processing was carried out to determine peak areas and retention times.

For all the samples a 0.5 µl syringe was used and 0.01 µl of sample was injected, except for samples run with methods A, B and C where a 1 µl syringe was used to inject 0.02 µl of sample. This was done because the 0.5 µl syringe was only received at a later stage. Injecting such small amounts is not ideal, but since only relative amounts were compared, the method was acceptable. In future more repeatable results could be obtained by injecting sufficiently small amounts directly on the column (splitless injection).

A split/splitless inlet was used and all samples were split in a 50:1 ratio. The inlet temperature was kept at 250 °C. Helium was used as the carrier gas at a flow rate of 1 ml/min. The only exception was for samples that were analysed using method C, where a carrier gas flow rate of 1.5 ml/min was used. All analyses were performed under corrected constant flow via pressure ramps (as recommended by the ChromaTOF® software for two-dimensional analyses). All the samples were analysed while an inlet liner containing deactivated glass wool was installed, except for the samples analysed using method J. This is commented on in the discussion of the results (see Chapter 6, Section 1).

Table 5.3 Summary of different GC oven and modulator parameters used for optimised separation of samples.

GC Method	Primary oven temperature programme			Secondary oven temperature programme			Modulator	
	Rate (°C/min)	Target Temp (°C)	Duration (min)	Rate (°C/min)	Target Temp (°C)	Duration (min)	Modulator temperature offset* (°C)	Modulation period (s)
A	Initial	40	0.2	Initial	55	0.2	30	4
	2	200	5	2	215	5		
B	Initial	40	0.2	Initial	55	0.2	30	4
	2	240	5	2	255	5		
C	Initial	40	0.2	Initial	60	0.2	35	6
	3	200	5	3	220	5		
D	Initial	40	0.2	Initial	60	0.2	35	4
	3	240	5	3	260	5		
E	Initial	45	0.2	Initial	50	0.2	20	2
	3	260	5	3	265	5		
F	Initial	45	0.2	Initial	55	0.2	25	3
	3	250	5	3	260	5		
G	Initial	40	0.2	Initial	60	0.2	35	4
	1.5	200	5	1.5	220	5		
H	Initial	40	0.2	Initial	60	0.2	35	5
	2	200	5	2	220	5		
I	Initial	40	0.2	Initial	60	0.2	35	5.4
	2	200	5	2	240	5		
J	Initial	10	10	Initial	50	0.2	95	6
	2	220	5	2.18	300	5		

* temperature relative to GC oven temperature, + 15 °C relative to secondary oven is recommended by ChromaTOF software.

The transfer line temperature was always higher than the final temperature of the second column (for reasons discussed in Chapter 5, Section 3.1). For GC methods A, B, C, D, G, H and I the transferline was set to 260 °C and for E and F the tranferline was set to 280 °C. A transferline temperature of 340 °C was used for method J.

The same mass spectrometer settings were used for all samples: the mass range was 30 – 500 amu. The lower limit was chosen in order to simplify the process of finding oxygenates, especially alcohols, which have a characteristic ion of mass 32. An acquisition rate of 100 spectra per second was used. The detector voltage was 1650 V and the electron energy was -70 V. The ion source temperature was set to 250 °C.

During data processing a signal-to-noise ratio (S/N) of 500 and a baseline offset of 1 (ChromaTOF® software defines this as “just above the noise”) was used. For trace analysis lower S/N ratios are usually recommended, but since classifications were done manually it was impractical to use lower S/N ratios for these complex samples.

A minimum match of 750 (out of 1000) was required to combine two peaks (as described in Chapter 3, Section 3.2). The minimum S/N of 6 was required for a subpeak to be retained. A traditional integration approach was used to determine the relative area for each peak. The recorded mass spectra were compared to the following NIST libraries using a forward method: mainlib, replib, nist_msms and nist_ri.

4.4 Analysis of samples

As mentioned previously, different methods were used to analyse the different samples. These methods are summarised in Table 5.4 below. The codes used in the table refer to the parameters and column sets in Tables 5.2 and 5.3 respectively.

Table 5.4 Summary of the parameters and column sets used to analyse the different samples.

Sample name	Column Set			
	1	2	3	4
I	A*	C*	E*	J
II	A*	C*	E*	J
III	B*	D*	F*	J
IV	B*	D*	F*	J
V	B*	D*	F*	J
VI	B*	D*	F*	J
VII	G*	H*	E*	J
VIII	G*	H*	E*	J
IX	H*	I*	F*	J

*These samples were analysed with an inlet liner containing glass wool.

The methods used for column sets 1, 2 and 3 were based on methods used in literature [8,9]. These methods were altered slightly as needed to optimise the separation of each individual sample (based on the aims set out in Chapter 5, Section 3). For column set 4, there was no literature available on the separation of diesel samples using the SLB-IL100 column in the 1D of GCxGC. Therefore the optimisation of separation using column set 4 was investigated in depth and will be discussed in detail below. Research has been done on biodiesel samples using a SLB-IL100 column although the focus was on separating fatty acid isomers [11]. GCxGC systems with other (less polar) IL columns in the 1D have been used for the separation of diesel fuels [13].

4.4.1 Optimisation of separation on SLB-IL100

In an attempt to optimise the separation method for the different diesel samples on column set 4, a mixture of all samples was analysed and the separation optimised (details will be discussed below). The mixture was prepared by mixing 200 μ l of each of the nine samples in a GC autosampler vial. Once an optimised separation method for the mixture was attained, the samples were analysed individually using this method.

Initially a temperature program similar to the other column sets was used where the difference in oven temperatures stayed constant throughout the analysis. It was observed that a lot of wrap-around occurred (especially for AIHC) and therefore a modulation period of six seconds was used in further analyses. A carrier gas flow rate of 1.5 ml/min was used initially in order to prevent wrap-around, but this did not have a considerable effect and therefore a carrier gas flow rate of 1 ml/min was later employed. All analyses were performed under corrected constant flow via pressure ramps. The same mass spectrometer settings were used as given in Chapter 5, Section 4.3. For comparative reasons, the same x-axis (retention time in ${}^1\text{D}$) for all the chromatograms are given in the figures below. Table 5.5 contains additional details of the temperatures used during optimisation procedures followed by a detailed discussion of the optimisation procedure. The separation was optimised in order to find the separation conditions where the chromatographic peaks are maximally distributed throughout the two-dimensional chromatogram (with minimal wrap-around) and compound classes are resolved from one another (see Chapter 5, Section 3.2).

Table 5.5 Additional details regarding the different temperature programs used during optimisation and corresponding figures.

First oven			Second oven			Corresponding figures
Initial temp ($^{\circ}\text{C}$) / hold (min)	Temp ramp ($^{\circ}\text{C}/\text{min}$)	Final temp ($^{\circ}\text{C}$) / hold (min)	Initial temp ($^{\circ}\text{C}$) / hold (min)	Temp ramp ($^{\circ}\text{C}/\text{min}$)	Final temp ($^{\circ}\text{C}$) / hold (min)	
40 / 1	2	200 / 5	70 / 1	1.75	210 / 5	5.3 (1 st attempt) and 5.4
40 / 1	2	200 / 5	90 / 1	1.5	210 / 5	5.3 (2 nd attempt) and 5.5
40 / 1	2	200 / 5	90 / 1	1.5	210 / 5	5.3 (2 nd attempt) and 5.6 *
40 / 10	2	200 / 5	70 / 10	1.75	210 / 5	5.7 and 5.8
30 / 20	2	200 / 5	70 / 0.2	1.5	230 / 5	5.9 (1 st attempt) and 5.10
30 / 20	2	200 / 5	90 / 0.2	1.34	230 / 5	5.9 (2 nd attempt) and 5.11
30 / 20	2	200 / 5	90 / 0.2	1.34	230 / 5	5.9 (2 nd attempt) and 5.12 #
10 / 10	2	220 / 5	50 / 0.2	2.18	300 / 5	5.13 and 5.14

* A longer modulation period was used

The pressfit connector was placed closer to the modulator

As explained, it was crucial that the polar fraction of the samples were resolved in terms of classes of compounds. In order to gain optimum separation conditions for the more retained compounds (polar species), the temperature difference between the ovens was decreased with time. The temperature program is shown in Figure 5.3 (1st attempt) and the resulting chromatogram is given in Figure 5.4.

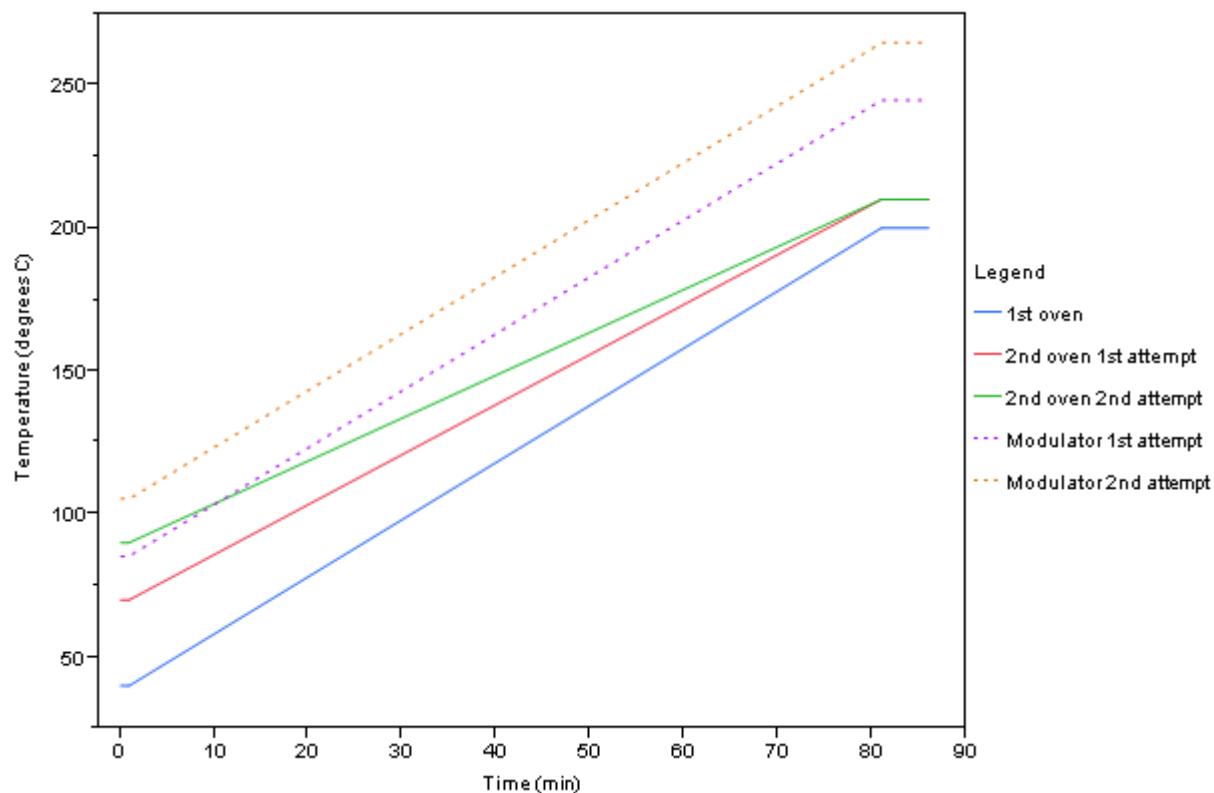


Figure 5.3 Two temperature programs (1st and 2nd attempt) used during the optimisation of separations on the SLB-IL 100 column in the ^1D of GCxGC.

The temperature program was then altered (Figure 5.3, 2nd attempt) by increasing the initial temperature difference to 50 °C (by increasing the initial temperature of the second oven). The modulator offset was adjusted accordingly. Figure 5.5 shows the resulting chromatogram. Wrap-around of AIHC occurred to a lesser extent under these conditions, but class resolution was not yet obtained. The polar species were not separated as completely compared to previous attempts (compare Figure 5.4).

To further eliminate wrap-around, the modulation period was increased to eight seconds, whilst temperature programming remained the same (see Figure 5.3, 2nd attempt). The resulting chromatogram is shown in Figure 5.6. Less wrap-around of AIHC did occur, but the ^1D peak widths were between 15 and 20 seconds wide, and

consequently each ^1D peak was not modulated 3 times as required (see Chapter 3, Section 3.1). For this reason a shorter modulation period and oven temperatures (especially temperature differences between the two ovens) was used to optimise future separations.

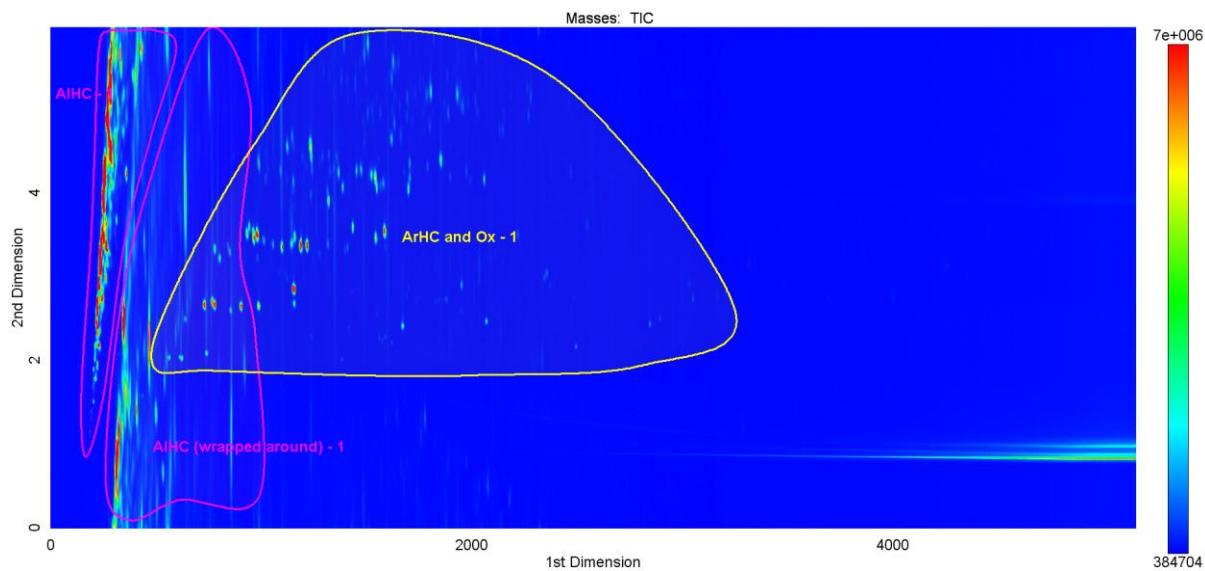


Figure 5.4 The chromatogram displaying the separation achieved when the temperature difference between the two ovens decreases from 30 °C to 10 °C. The categories shown are AIHC (—) as well as ArHC and Ox (—) which elute in the same region.

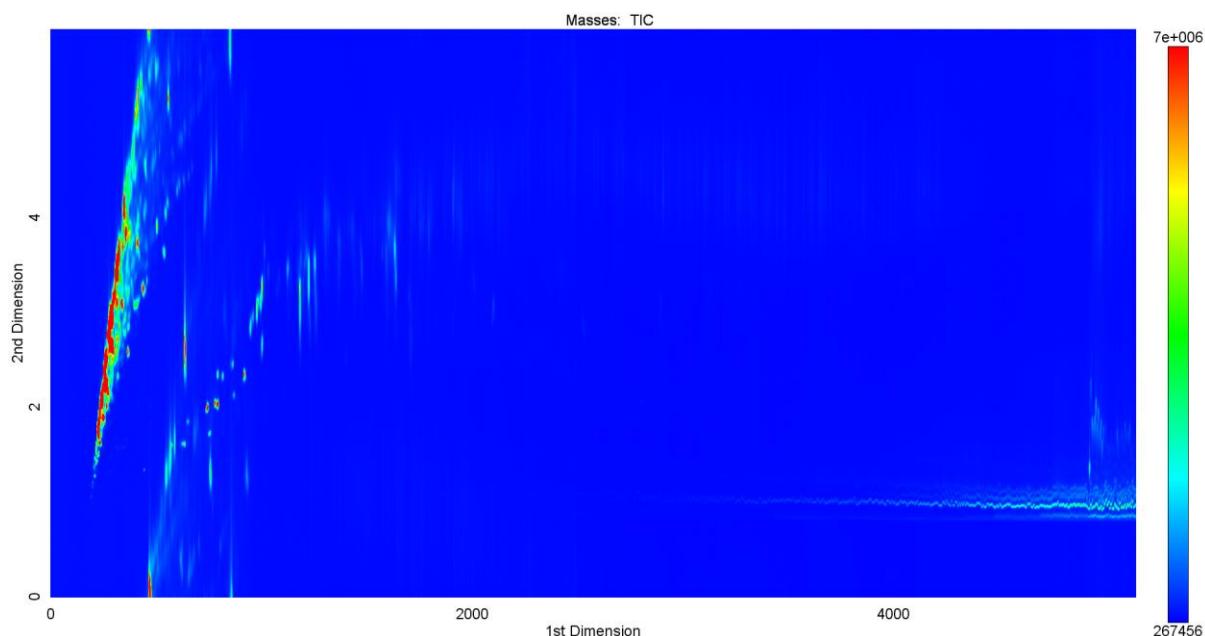


Figure 5.5 The chromatogram displaying the separation achieved when the temperature difference between the two ovens decreases from 50 °C to 10 °C with a modulation period of 6 seconds.

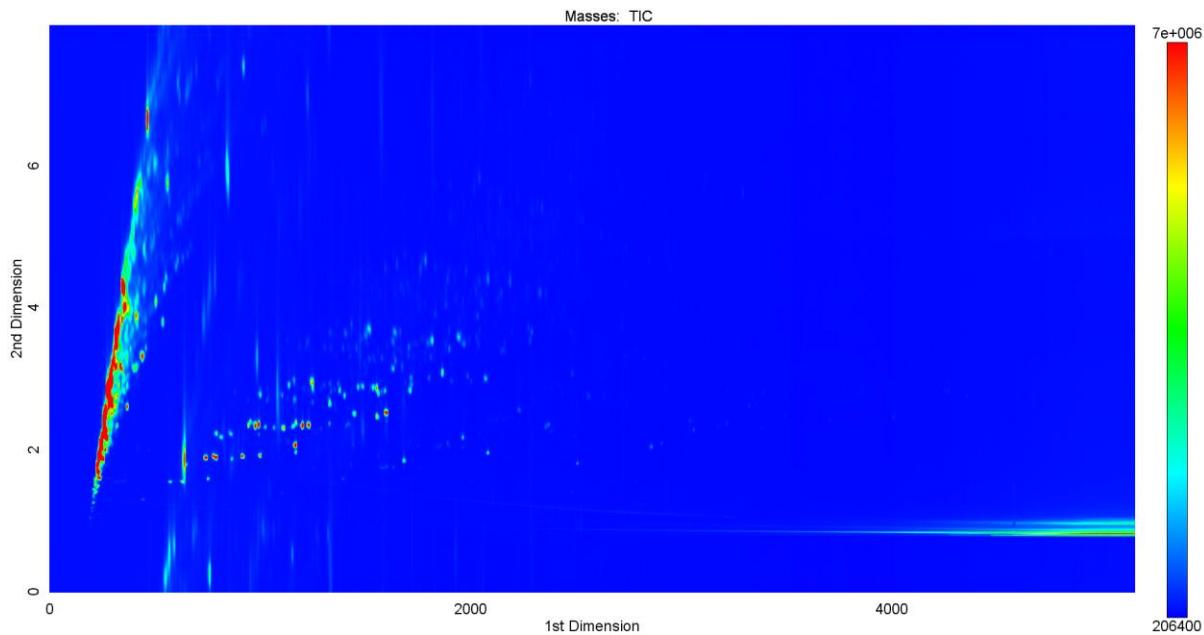


Figure 5.6 The chromatogram displaying the separation achieved when the temperature difference between the two ovens decreases from 50 °C to 10 °C with a modulation period of 8 seconds.

It was decided to use isothermal conditions for the initial separation of AIHC (which elute at lower temperatures) as shown in Figure 5.7. This was done in an attempt to allow only the AIHC and smaller polar species to enter the second column initially, after which the remaining (more retained) polar species can be separated under more ideal conditions (smaller temperature difference between ovens). Extreme wrap-around of AIHC occurred, but the polar species were better separated (Figure 5.8).

At this stage it became clear that two separate analyses were necessary to obtain an optimised separation. One separation could be optimised for the separation of the polar species (using the method just described gives excellent distribution of these compounds and compound classes), whilst the other could focus on the separation of AIHC (with little or no wrap-around). For practical reasons two different separations for each sample was not considered, since the analysis time will effectively be doubled. If such a long analysis time can be allowed, an alternative approach can be used to separate the sample in one analysis. In order to eliminate the wrap-around it is common practice to increase the temperature ramp in the ²D. In this specific case this will not be beneficial, since this will result in the polar species not being separated sufficiently in the ²D (the temperature will be higher) and these

compounds may even be unretained. To circumvent this situation it would be more beneficial to decrease the temperature ramp in the ^1D and ^2D to allow an increase in the modulation period (without reduction in the number of ^2D chromatograms per ^1D peak). A longer ^2D analysis time will result in less wrap-around of AIHC. Note that the peaks will broaden when using a slower temperature ramp, but this will not influence resolution. This was not investigated due to total analysis time considerations.

It was not possible to achieve ideal separation of polar and non-polar compounds in a single analysis without increasing the analysis time and a compromise had to be made. In subsequent optimisation procedures attempts were made to minimise wrap-around of non-polar analytes as much as possible whilst avoiding considerable loss of class resolution of specifically the more polar analytes.

During a single modulation period, both (more volatile) polar and (less volatile) non-polar analytes are injected into the second column. The boiling point distribution of this fraction is unusually large. This unique situation is a direct result of the high polarity of the IL column in the ^1D . The system is highly orthogonal compared to more conventional systems, since separation in the ^1D is based mainly on polar interactions and not vapour pressures. The nature of the second column does not allow for significant retention of polar species (low k -values, resulting in poor resolution) but the non-polar species are retained to a large extent. Consequently the general elution problem is encountered (see Chapter 3, Section 2.2) since the ^2D separation occurs under essentially isothermal conditions. Accordingly wrap-around of non-polar analytes was unavoidable since sufficient separation of polar analytes was a main concern. Wrap-around was tolerated since the wrapped-around compounds could be distinguished from the polar species based on their ^2D peak widths and mass spectra. This approach would not have been successful if only FID was used as detection method since peak deconvolution (which will be necessary in regions where polar analytes and wrapped-around compounds overlap) is based on mass spectral data.

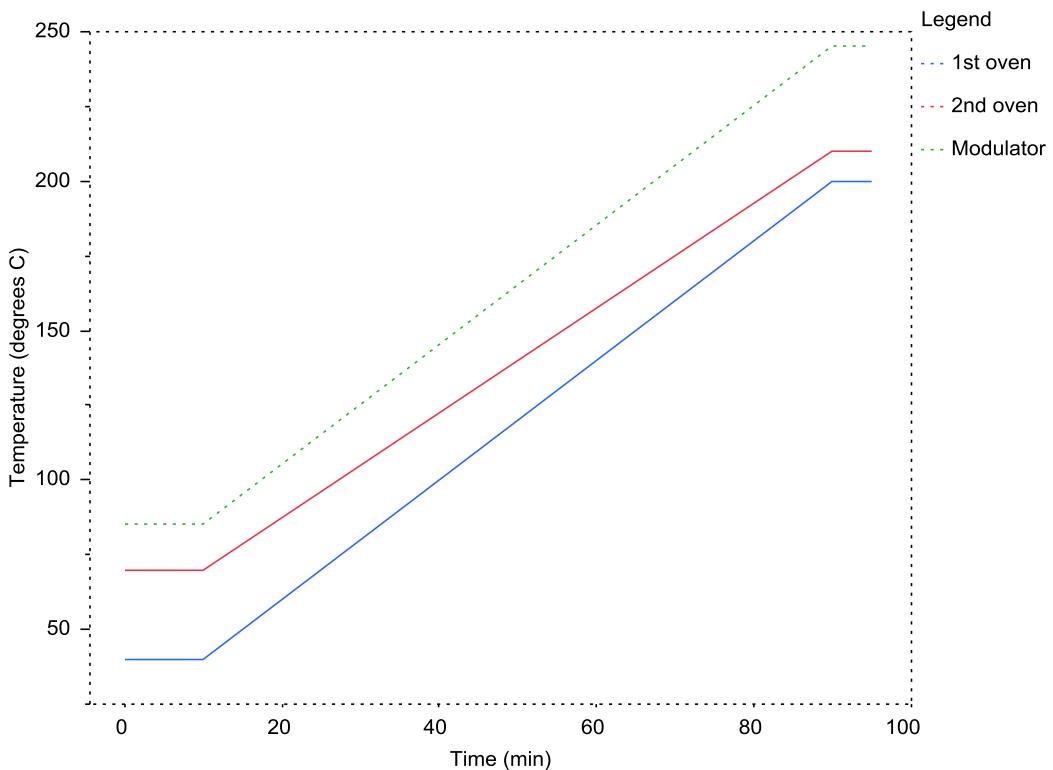


Figure 5.7 An alternative temperature programs used during the optimisation of separations on the SLB-IL 100 column in the ^1D of GCxGC.

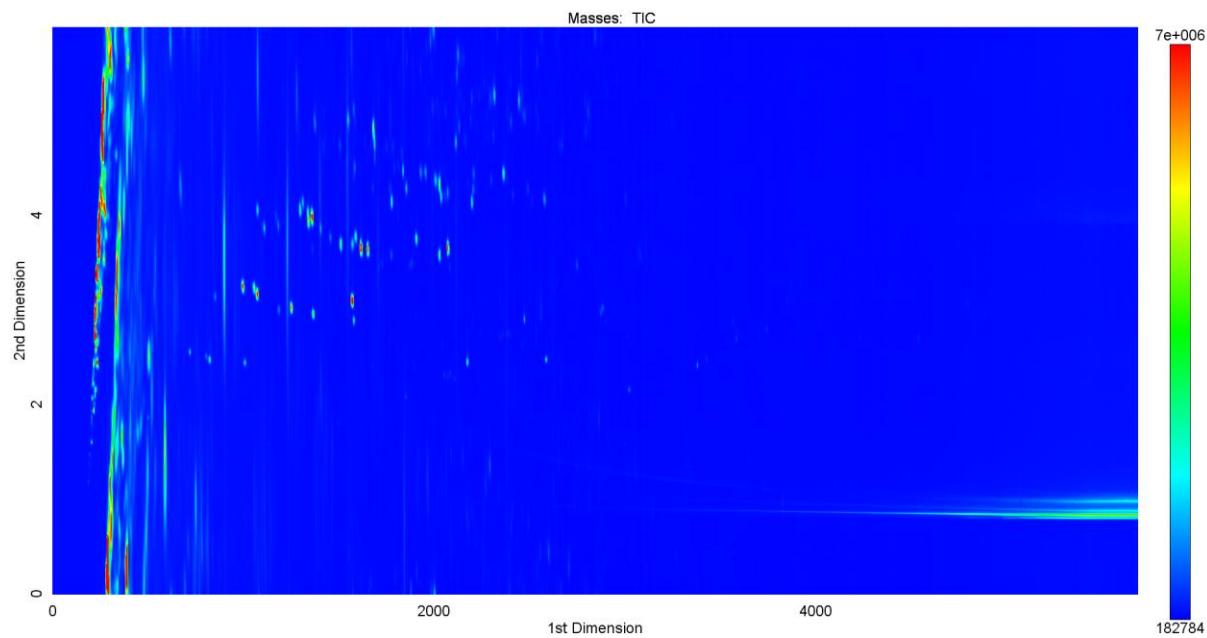


Figure 5.8 The chromatogram displaying the separation achieved when isothermal conditions were used initially.

To improve the separation of AIHC, it was decided to initially increase the second oven temperature whilst the first oven temperature stayed constant. The initial temperature of the first oven was also lowered. The temperature program is shown in Figure 5.9 (1st attempt) and the resulting chromatogram in Figure 5.10. Separation

of AIHC seemed to improve, whilst sufficient separation of polar species was still obtained (although the classes of compounds are much closer to one another compared to the separation achieved in Figure 5.8).

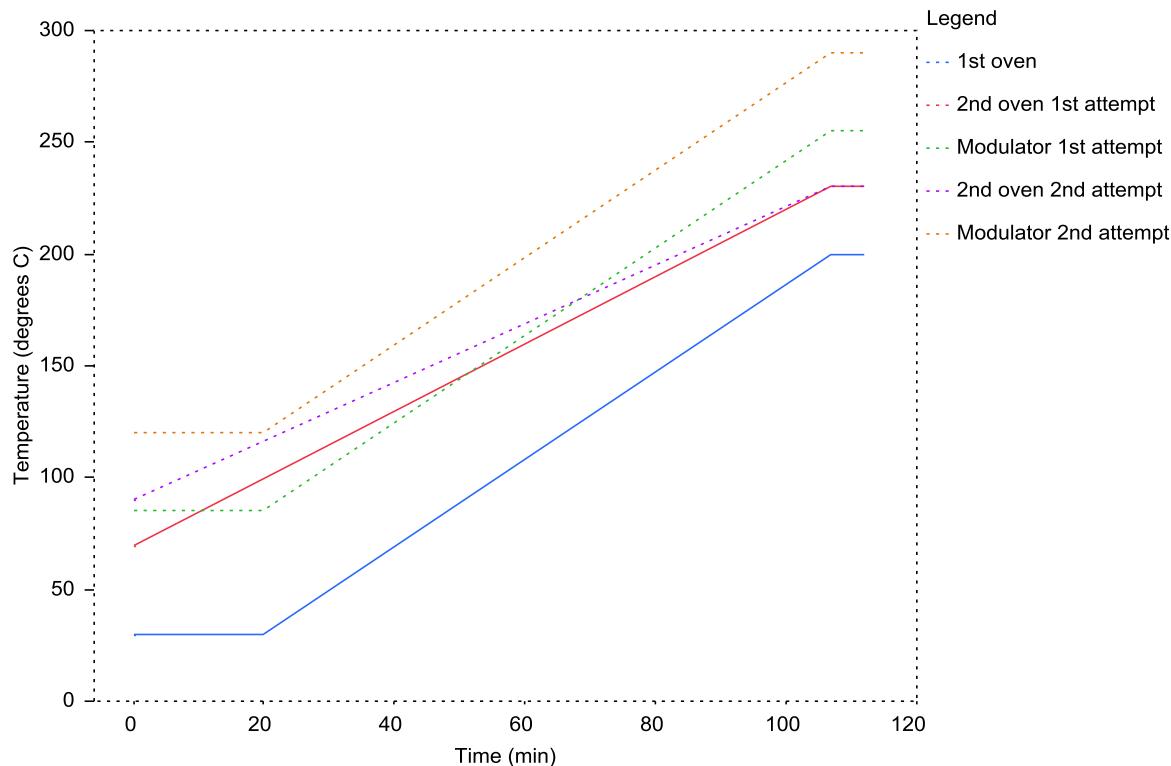


Figure 5.9 Two alternative temperature programs (1st and 2nd attempt) used during the optimisation of separations on the SLB-IL 100 column in the ¹D of GCxGC.

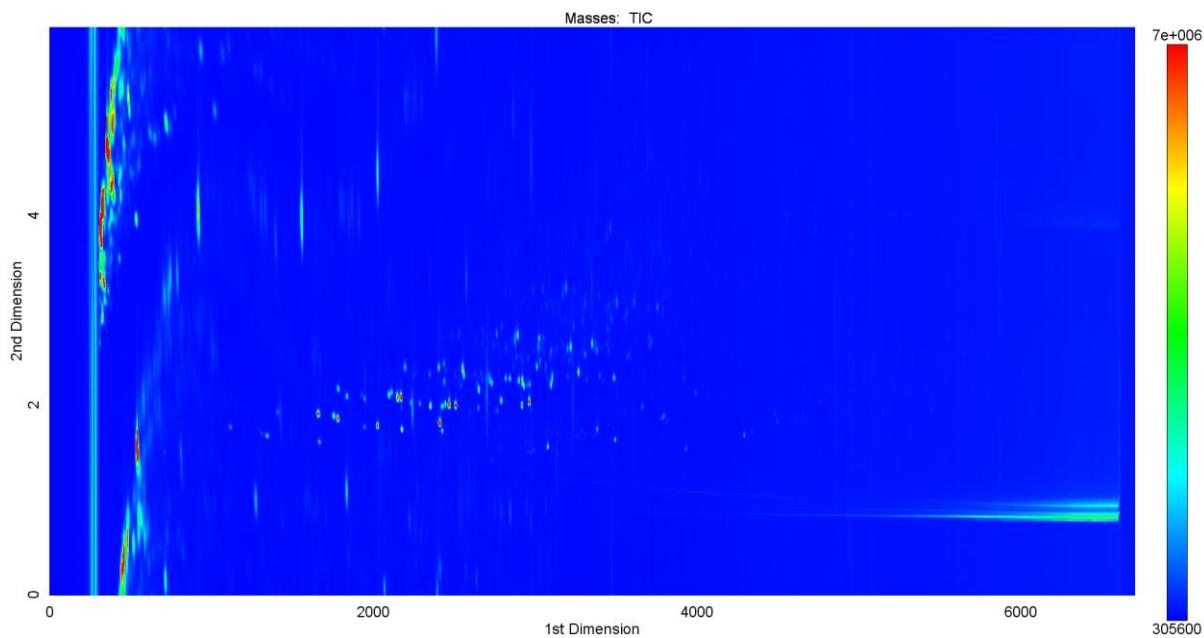


Figure 5.10 The chromatogram displaying the separation achieved when only the first oven was initially kept isothermal. The initial temperature difference was 40 °C.

From previous attempts it was clear that a larger initial temperature difference (between the two ovens) decreased wrap-around. The initial temperature of the second column was again increased to give a larger initial temperature difference for the two ovens (Figure 5.9, 2nd attempt). The resulting chromatogram is given in Figure 5.11. The initial wrap-around was reduced and classes of polar species could still readily be distinguished.

As shown in Figure 3.4, the two columns are connected with a pressfit connector before the modulator. The amount of the second column outside of the modulator proved to play an important role in facilitating wrap-around. The non-volatile AIHC (high molecular mass) are not retained on the IL column (¹D) but are then retained strongly in the part of the second column just before the modulator. Thus the distance of the second column between the pressfit connector and the modulator had to be minimized. The effect of this can be seen by comparing Figure 5.11 and 5.12 where the same separation conditions were used, but the distance of the second column between the pressfit connector and the modulator was changed. In Figure 5.11 this distance was about 5 cm (this is usually done to conveniently change the first dimension column as necessary without unwinding the second column). In Figure 5.12 this distance was reduced to only \pm 5 mm. Care should be taken when placing the pressfit connector very close to the modulator, since column strain increases and the column can easily break while in this position.

For further separation of the (more volatile) AIHC, it was necessary for sub-ambient temperatures in the initial stages of the temperature program for the first oven. At these lower temperatures, the less volatile AIHC will elute later from the first column, allowing time for the more volatile compounds to be separated first. By lowering the initial temperature of the first oven to 10 °C (Figure 5.13), and by keeping the initial temperature difference at 40 °C (as before), sufficient separation of both AIHC and more polar species were obtained as shown in Figure 5.14. The final temperature difference between the two ovens was also increased to minimise wrap-around of AIHC. These conditions were used to analyse all nine samples individually (see Table 5.3 and 5.4).

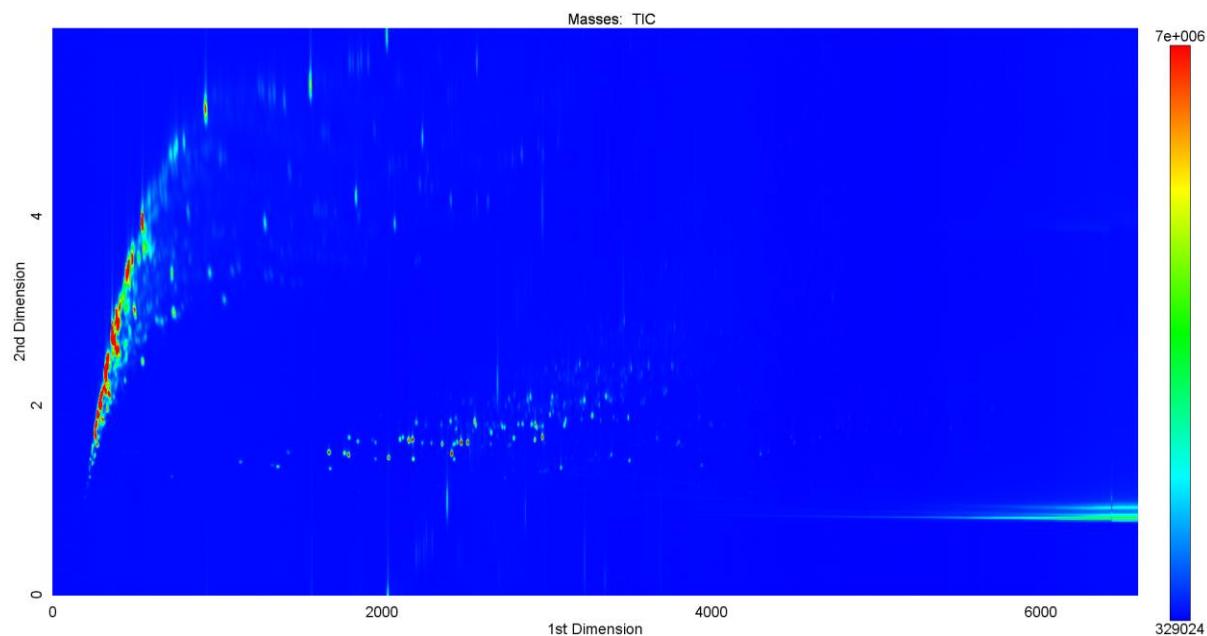


Figure 5.11 The chromatogram displaying the separation achieved when only the first oven was initially kept isothermal. The initial temperature difference was 60 °C. The length of column between the pressfit connector and the modulator was relatively large (± 5 cm).

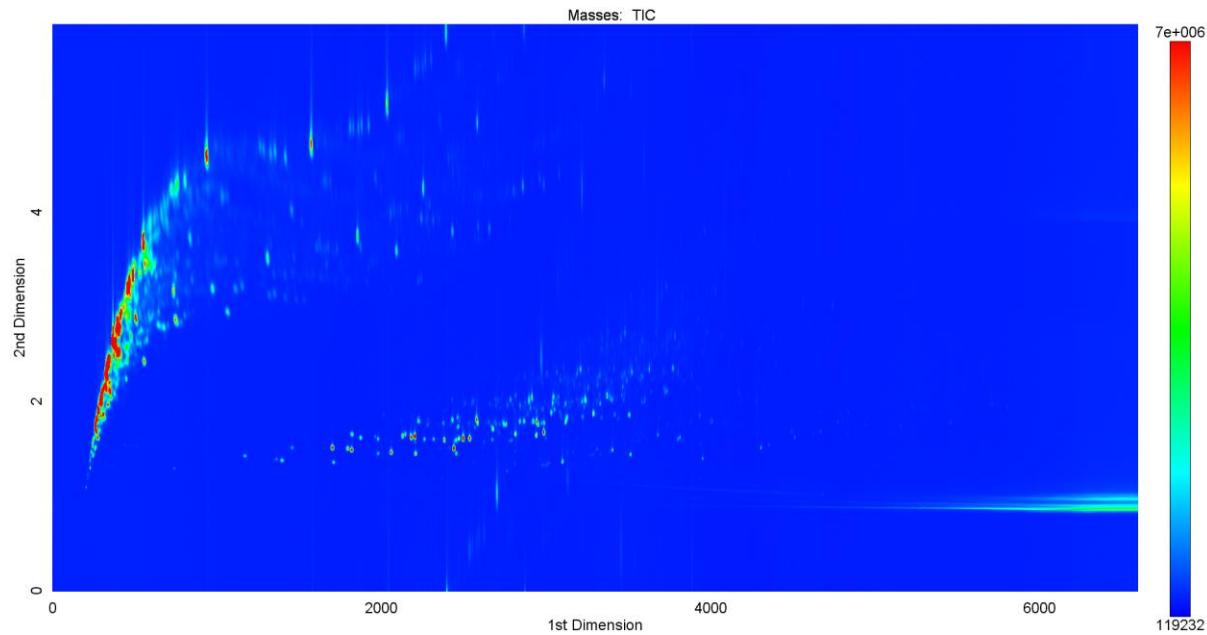


Figure 5.12 The chromatogram displaying the separation achieved when only the first oven was initially kept isothermal. The initial temperature difference was 60 °C. The length of column between the pressfit connector and the modulator was very small (± 5 mm).

Separation of the more polar species (ArHC and OX) was not ideal since the rooftiles are not as easy to distinguish from one another (compared to Figures 5.8 and 5.9). The remaining wrap-around of the non-polar analytes was also not ideal. This is a direct result of the high orthogonality of the system as explained previously.

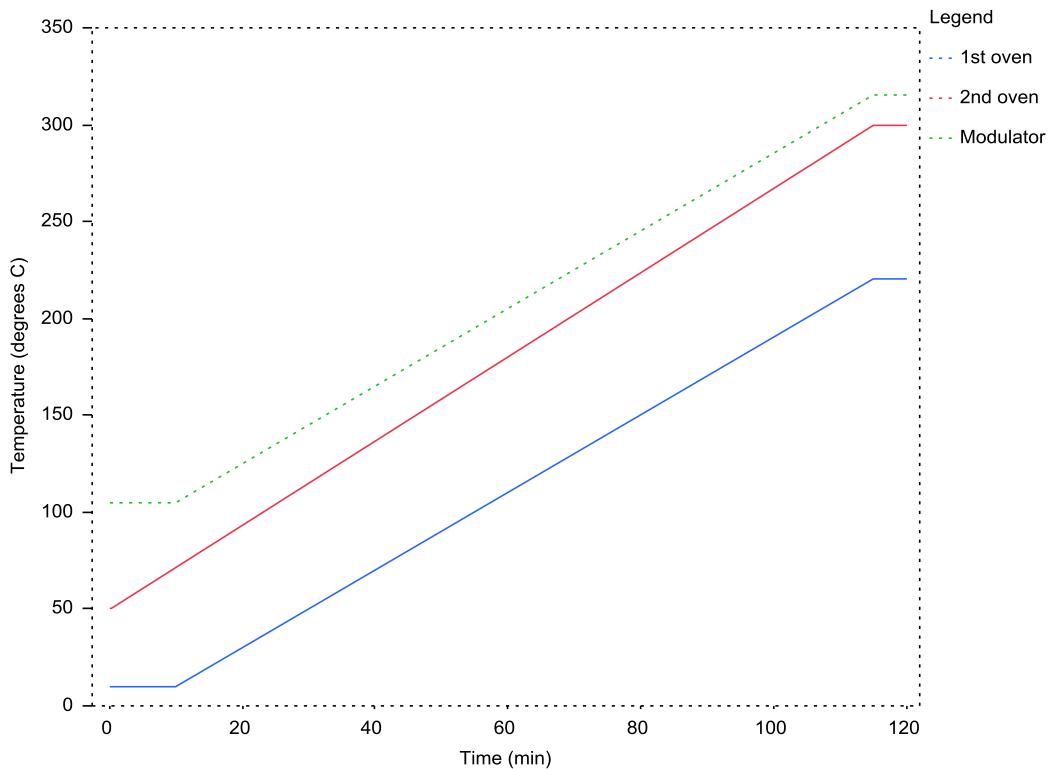


Figure 5.13 An alternative temperature programs used during the optimisation of separations on the SLB-IL 100 column in the ^1D of GCxGC.

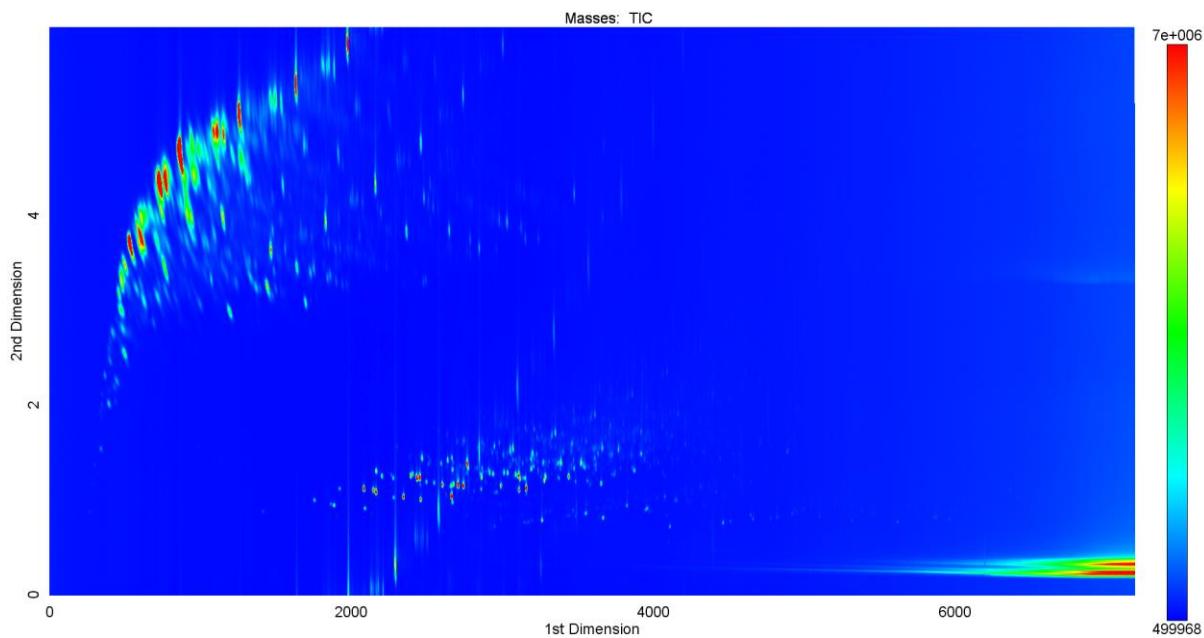


Figure 5.14 The chromatogram displaying the separation achieved for the mixture of all the samples under optimised conditions.

5. Statistical Analyses

Once the compositional data was available, it was compared with the corresponding measured WSD value for each sample. Initially the main categories (AIHC, ArHC and Ox) were considered, and a mixture model approach (see Chapter 4, Section 6) was followed. This agrees with what has been done in literature when determining a model for the physico-chemical relationship for jet and diesel fuels [14]. Several physical properties were investigated including cetane index, smoke point, freezing point and cloud point [14].

An alternative chemometric approach to calculate diesel fuel characteristics has been proposed in literature [15] and entails cluster analysis (factor analysis), regression analysis and principal component analysis. The specific consumption of fuel and the exhaust gas smoke of 92 fuel samples were investigated [15]. It was not possible to follow the latter approach due to the limited number of samples available. Linear multiple regression (see Chapter 4, Section 4) could not be used since the independent variables are dependent on one another (compositional data imply a constant total amount).

To observe the effect of chain length, ternary plots were used to determine trends in WSD based on longer and shorter chains. Compounds with chains containing twelve or more carbon atoms (≥ 12) are often used as lubricity additives (see Chapter 2, Section 3.3) [16] and therefore shorter chain compounds (< 12) are expected to contribute less towards improving diesel lubricity. Linear associations between individual compounds and WSD were also investigated to get an indication of the effect of individual compounds and lubricity. To confirm that the highly correlated compounds strongly influence lubricity, further experimental proof is needed and thus results were seen only as an indication of which compounds may be important.

All statistical calculations were performed using JMP[®] 8.0.1, a statistical software package from SAS[®] Institute Inc. (Cary, North Carolina, USA). This software package was chosen above other commercially available packages because of its availability and user-friendliness as well as the availability of support by professional

statisticians. A week-long training course regarding this software was attended where an overview of the programme functions and features were given. A professionally written user-guide was also supplied and used frequently during initial stages of statistical analysis [17].

References

1. Ryan, D.; Morrison, P.; Marriott, P. *Journal of Chromatography A* **2005**, 1071, pp. 47 – 53.
2. Dalluge, J.; Vreuls, R.J.J.; Beens, J.; Brinkman, U.A.Th. *Journal of Separation Science* **2002**, 25, pp. 201 – 214.
3. Beens, J.; Blomberg, J.; Schoenmakers, P.J. *Journal of High Resolution Chromatography* **2000**, 23, pp. 182 – 188.
4. Ong, R.; Marriott, P.; Morrison, P.; Haglund, P. *Journal of Chromatography A* **2002**, 962, pp. 135 – 152.
5. Beens, J.; Boelens, H.; Tijssen, R.; Blomberg, J. *Journal of High Resolution Chromatography* **1998**, 21, pp. 47 – 54.
6. Vendeuvre, C.; Ruiz-Gerrero, R.; Bertoncini, F.; Duval, L.; Thiébaut, D. *Oil & Gas Science and Technology – Revue d’ IFP* **2007**, 62, pp. 43 – 55.
7. http://www.leco.com/products/sep_sci/pegasus_4d/pegasus_4d.htm
8. Van der Westhuizen, R.; Crouch, R.; Sandra, P. *Journal of Separation Science* **2008**, 31, pp. 3423 – 3428.
9. Adachour, M.; Beens, J.; Vreuls, R.J.J.; Batenburg, A.M.; Brinkman, U.A.Th. *Journal of Chromatography A* **2004**, 1054, pp. 47 – 55.
10. Peters, D.G.; Hayes, J.M.; Hieftje, G.M. *Chemical separations and measurements: The theory and practice of analytical chemistry*; W.B. Saunders Company: United States of America, **1974**, pp. 568 – 571.
11. Ragonese, C.; Tranchida, P.Q.; Sciarrone, D.; Mondello, L. *Journal of Chromatography A* **2009**, 1216, pp. 8992 – 8997.
12. Armstrong, D.W.; Payagala, T.; Sidisky, L.M. *LC-GC Europe* **2009**, 22, pp. 459 – 467.
13. Seeley, J.V.; Seeley, S.K.; Libby, E.K.; Breitbach, Z.S.; Armstrong, D.W. *Analytical and Bioanalytical Chemistry* **2008**, 390, pp. 323 – 332.
14. Cookson, D.J.; Smith, B.E. *Energy & Fuels* **1992**, 6, pp. 581 – 585.
15. Khots, M.S.; Nazarov, V.I.; Rudyak, K.B. *Chemometrics and Intelligent Laboratory Systems* **1994**, 22, pp. 265 – 271.

16. Mang, T.; Dresel, W. *Lubricants and Lubrication - second, completely revised and extended edition*; Wiley-VCH: Chichester, **2007**, pp. 107 – 113.
17. Sall, J.; Creighton, L.; Lehman, A. *JMP® Start Statistics: A Guide to Statistics and Data Analysis Using JMP® – 4th Edition*; SAS Institute Inc.: North Carolina, **2007**.

Chapter 6: GCxGC-TOFMS Results

Contents

1.	GCxGC-TOFMS results and data handling	90
2.	Comparison of different column sets	92
2.1	Graphical comparison	98
2.2	Mathematical comparison	106
3.	Conclusion.....	109
	References.....	110

1. GCxGC-TOFMS results and data handling

The two-dimensional chromatograms obtained for the different samples, using the different column sets, are given in Appendix A. After separation of the fuels, the chromatographic peaks were classified as described in Chapter 5, Section 3.3.

All the samples were classified in as much detail as possible for a signal-to-noise ratio of 500 or higher. Lower signal-to-noise ratios were not considered due to the complexity of the samples: up to 6000 compounds were detected with a S/N of 500 for some samples. Note that individual compounds may have S/N of less than 500, but as a class of compounds they may contribute significantly to the sample composition. Thus the results were not necessarily indicative of (mass) percentage composition, but relative compositions that can be compared to one another were obtained.

It could also be possible for compounds below this level to contribute to lubricity, but it was not possible to simultaneously characterise trace compounds together with the bulk of the mixture. In future it might be beneficial to compare the relationship between lubricity and specific compounds based on distinctive ion fragments present in their mass spectra (e.g. 31, 58, 60 and 85 for oxygenates). Internal standards can also then be used to calibrate the amounts present. Alternatively algorithms can be used to classify compounds based on their mass spectra rather than doing classifications manually.

Automated peak integration was performed on all samples to obtain the relative area percentage of each peak. Since similar compositions were obtained for the same sample analysed on the four different column sets (deviations will be discussed in Chapter 6, Section 2), this method of obtaining semi-quantitative results was considered sufficient.

The relative area percentage of each peak was used to give an indication of the amount of the particular compound present. These values (of peaks with S/N above 500) were normalised so that the sum of the area of all the sample peaks are 100 %.

Since very small amounts of the samples were injected, the repeatability of the exact volume injected could be influenced by various factors, and for this reason it is better to compare relative area percentages rather than absolute areas. Concerns regarding the injection method have been addressed in Chapter 5, Section 4.3. The column bleed was also integrated during data processing and differed depending on the columns used. By normalising sample peaks, the contribution of column bleed was discarded. The normalised values of peaks that were classified together (in the same class) were summed to calculate the relative percentage of the sample that is represented by each class. The class composition determined for each sample cannot be shown here for proprietary reasons, but it was reported to Dr S de Goede in a confidential list of tables in April 2011 (Sasol report 042011).

Classes of compounds were grouped together in subcategories and categories to simplify initial data analysis. The relative area percentages of all the classes in a group were summed to get the composition of each sample in terms of categories. These values are reported in Table 6.1. The different categories are AIHC, ArHC, and Ox (see Chapter 5, Section 3.3). By doing this, a lot of information gained from GCxGC analysis was lost. During statistical analysis, the number of observations (number of sample streams) must exceed the number of independent variables (number of separated classes). Since only nine samples were available, statistical analyses could not be based on classes of compounds (some samples contained more than 100 classes) and the composition of samples were only considered in terms of categories (only three per sample).

In cases where a '0'-value is reported, it is indicative that the specified compound class or category was not detected at a signal-to-noise ratio of 500. At lower levels these compounds may be present.

During the classification of samples, several challenges were experienced. For the first three column sets (see Table 5.2) it was difficult to distinguish between alkenes and cyclic alkanes, since these compounds elute together in both dimensions and have very similar mass spectra. Literature indicates that oxygen and nitrogen containing compounds as well as aromatic hydrocarbons play a more important role in fuel lubricity compared to other hydrocarbons (see Chapter 2 and references

therein). Consequently it was assumed that crucial information would not be lost if alkenes and cyclic alkanes were classified together as one class. Similar arguments hold for compounds with higher degrees of unsaturation (number of ring structures or π -bonds), which were also classified together.

The degree of unsaturation (also known as the “hydrogen deficiency index” - HDI) can be calculated from the molecular formula ($C_xH_yO_z$) using Equation 6.1. Oxygen atoms do not contribute to the hydrogen deficiency index. For certain column sets it was also difficult to distinguish between aldehydes and ketones and these compounds were also classified together.

$$HDI = x - y/2 + 1 \quad (6.1)$$

Table 6.1 Sample composition in terms of categories for each of the different column sets.

Column set	Category	I	II	III	IV	V	VI	VII	VIII	IX
1	AIHC	99.8	99.9	76.6	73.4	61.3	57.5	84.9	85.7	72.5
	ArHC	0.2	0.1	23.4	26.5	26.7	28.5	15.1	14.3	27.5
	Ox	0.0	0.0	0.0	0.1	12.0	14.0	0.0	0.0	0.0
2	AIHC	99.7	99.8	64.2	65.5	52.0	50.1	86.6	87.2	68.1
	ArHC	0.3	0.2	34.9	34.3	32.4	30.2	13.4	12.8	31.9
	Ox	0.0	0.0	0.9	0.2	15.7	19.4	0.0	0.0	0.0
3	AIHC	99.8	99.9	75.9	76.4	67.3	66.4	84.0	84.4	73.7
	ArHC	0.2	0.1	24.1	23.6	24.6	23.7	16.0	15.6	26.3
	Ox	0.0	0.0	0.0	0.0	8.1	9.9	0.0	0.0	0.0
4	AIHC	99.7	99.9	76.3	69.0	62.6	60.6	88.4	90.7	71.3
	ArHC	0.3	0.1	23.7	31.0	29.4	28.3	11.6	9.3	28.7
	Ox	0.0	0.0	0.0	0.0	8.0	11.1	0.0	0.0	0.0

2. Comparison of different column sets

The sample composition results obtained for the different column sets (shown in Table 6.1) differed slightly, and further investigation was needed. To compare these

results, ternary plots were used to show variations in the results obtained for each sample analysed on the four different column sets (see Figure 6.1).

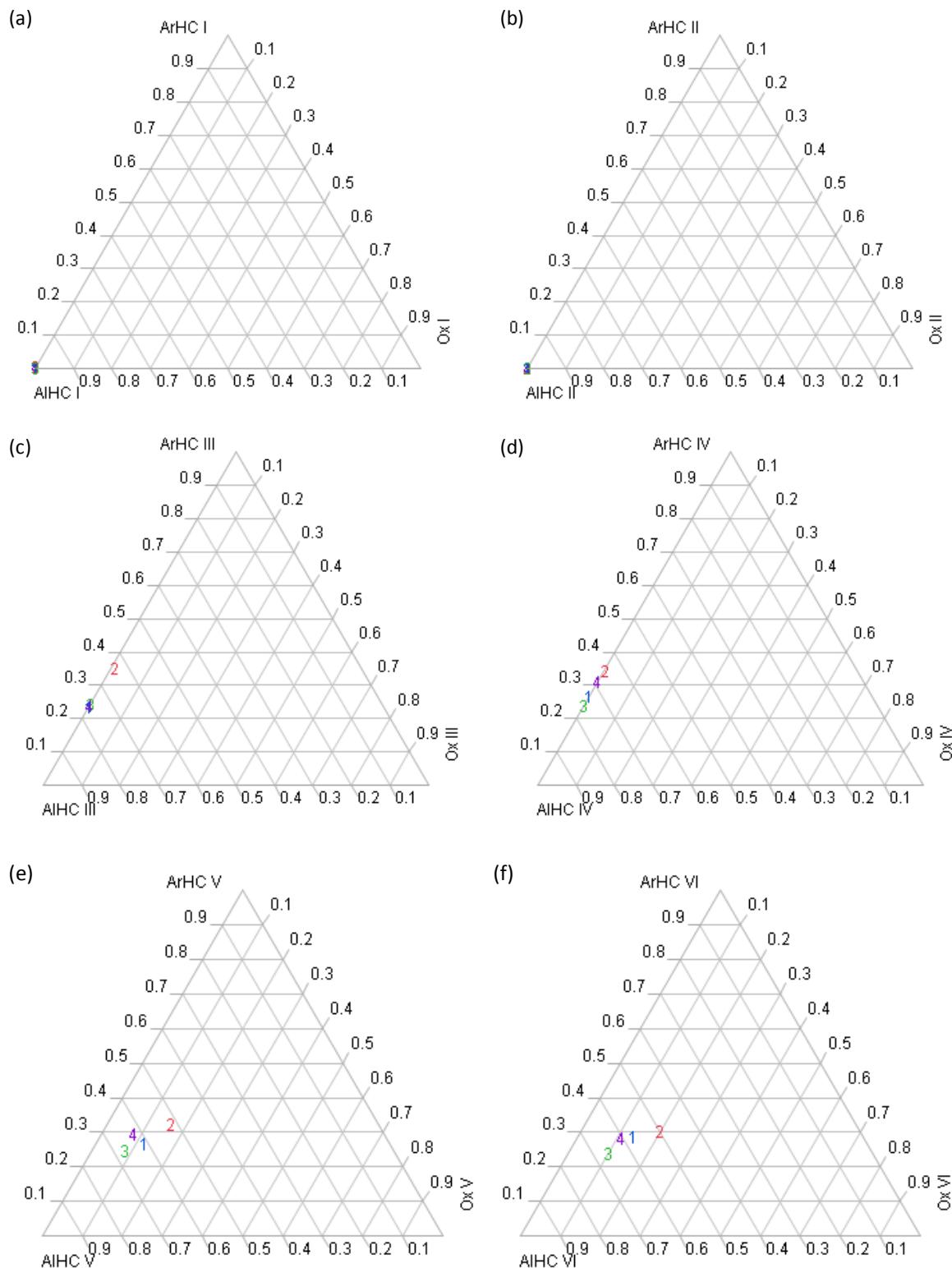


Figure 6.1 Continued on the next page.

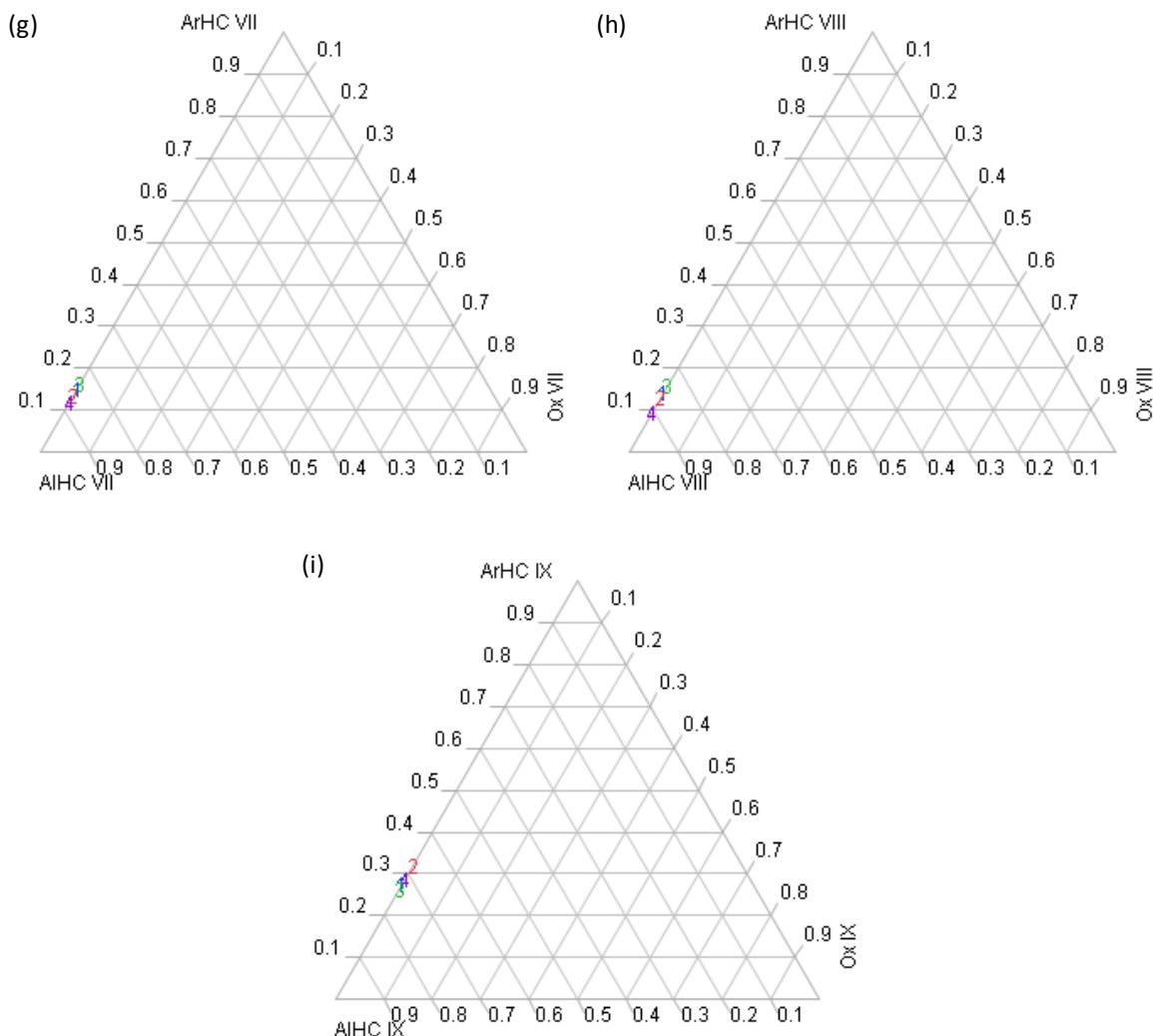


Figure 6.1 (Continued) Compositional data obtained with the four different column sets. The composition of samples I (a), II (b), III (c), IV (d), V (e), VI (f), VII (g), VIII (h) and IX (i) are shown. The number in each diagram indicates which column set (1, 2, 3 or 4) was used, and its position in the triangular space indicates the composition obtained for that specific column set. The horizontal axis represents the total amount of aliphatic hydrocarbons (AIHC) present. The axis on the left represents the total amount of aromatic hydrocarbons (ArHC) whilst the axis on the right represents the total amount of oxygenates (Ox) present.

The results show that similar compositions were obtained for samples I, II, VII, VIII, and IX when analysed on the four different column sets. It is interesting to note that these samples do not contain significant amounts of oxygenates. For samples III and to a lesser extent for sample IV, analysis on column set 2 showed that more ArHC (and consequently less AIHC) were present compared to the other column sets. When considering samples V and VI, it seems that more Ox and ArHC were detected in column set 2, whilst the compositions obtained with the other column sets did not differ considerably.

The presence or absence of glass wool in the inlet liner did not seem to affect the results drastically. No glass wool was used for the samples analysed on column set 4, and these samples do not show the presence of significantly larger amounts of polar species compared to the other column sets. Therefore it can be deduced that negligible amounts of polar species were lost due to adsorption onto the glass wool present in the inlet liner. It is thus proposed that glass wool be used in the inlet liner to assist in proper mixing and complete aerosol evaporation during a split injection as discussed previously (see Chapter 5, Section 3.1.4).

Figure 6.1 indicates that column set 2 is more sensitive for detecting Ox, and may also detect ArHC better than the other column sets. This can be explained when the different S in the 1D of the respective column sets are considered. Oxygenates, such as alcohols and carboxylic acids, are expected to tail excessively in the 1D of column sets 1, 3 and 4, as well as in the 2D of all the column sets (see Chapter 3, Section 2.4). Figure 6.2 shows the extracted ion chromatograms (EIC) for carboxylic acids detected on the different column sets.

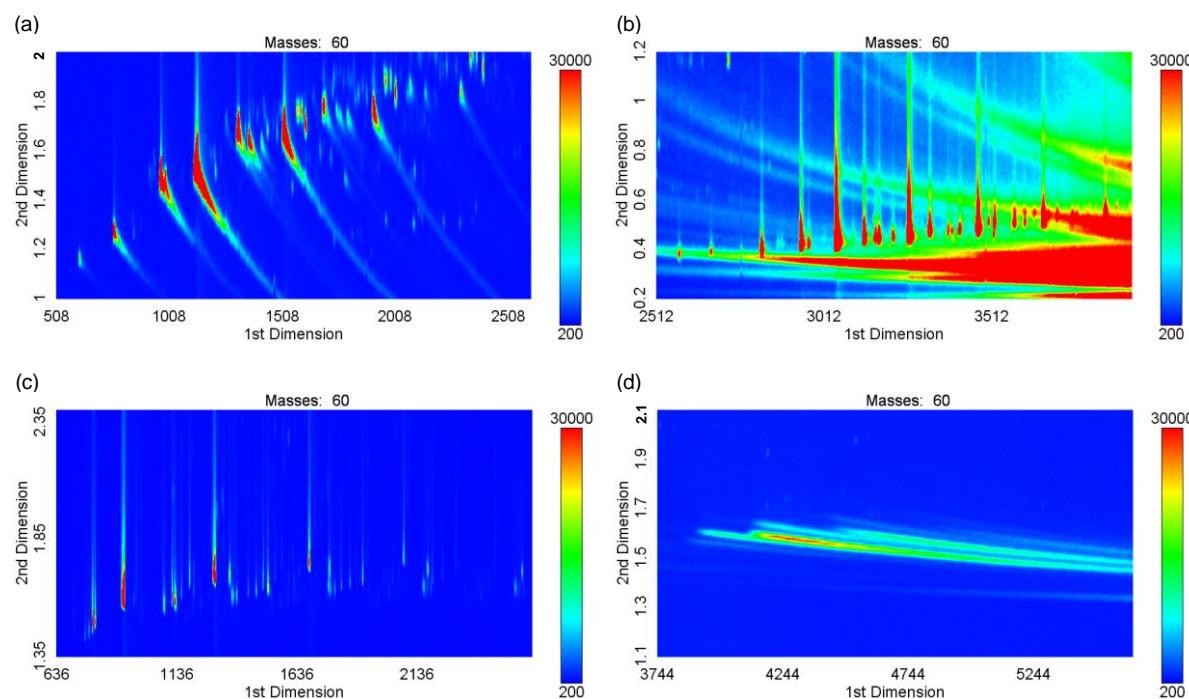


Figure 6.2 EIC of mass 60, indicating whether peak tailing was observed for carboxylic acids, analysed with column sets 1 (a), 2 (b), 3 (c) and 4 (d).

Peak tailing of carboxylic acids was not observed in the ^1D of column set 2, since analyte molecules would preferentially adsorb onto the end-groups of the PEG stationary phase and thus the silanol groups of the column walls are effectively deactivated. This has been discussed in detail (see Chapter 3, Section 2.4). As expected, peak tailing of carboxylic acids is observed in the ^1D of column sets 1 and 4 but not for column set 3. For column set 3 no peak tailing is observed, possibly because the glass walls of the ZB-5 column was sufficiently deactivated, and thus silanol groups are not exposed to the analytes. This corresponds with the ^2D of column sets 1, 2 and 4 where only slight peak tailing is observed in this dimension (a very similar column, ZB-5MSi was used in the ^2D of these column sets).

In column set 4 the IL stationary phase (^1D column) was expected to give peak tailing for highly polar compounds (see Chapter 5, Section 4.2) [1]. The reason for this tailing may be due to the ionic nature of the S, which could deprotonate the carboxylic acid and thus it would exist in an ionic form when in S. Thus the peak tailing was due to a kinetic effect, where the ionic form had to be protonated again before being able to partition into M.

For column set 1 peak tailing is observed in the ^1D , suggesting less effective silanol deactivation occurred in the Rxi-17 column, causing strong interactions with the highly polar carboxylic acids. This corresponds with the peak tailing observed in the ^2D of column set 3, where the same type of column was used (see Chapter 5, Table 5.2).

Column set 2 (PEG stationary phase in ^1D) detected more ArHC and less AIHC than the other column sets, which cannot be explained by peak tailing. This was observed for samples III, IV, V and VI. To determine whether the column set is more sensitive towards the detection of ArHC or less sensitive towards AIHC, the distribution of compounds (according to chain length) were investigated as shown in Figure 6.3.

When comparing these results only the variation in terms of chain length was considered, since variation in the relative area percentage (area %) for each column set is dependent on the amount of other components present: the relative amounts of AIHC, ArHC and Ox present are interdependent and therefore caution should be

taken when making conclusions based on the relative area percentages. The same compounds are expected to be detected for the same sample analysed on different column sets and thus this was investigated further.

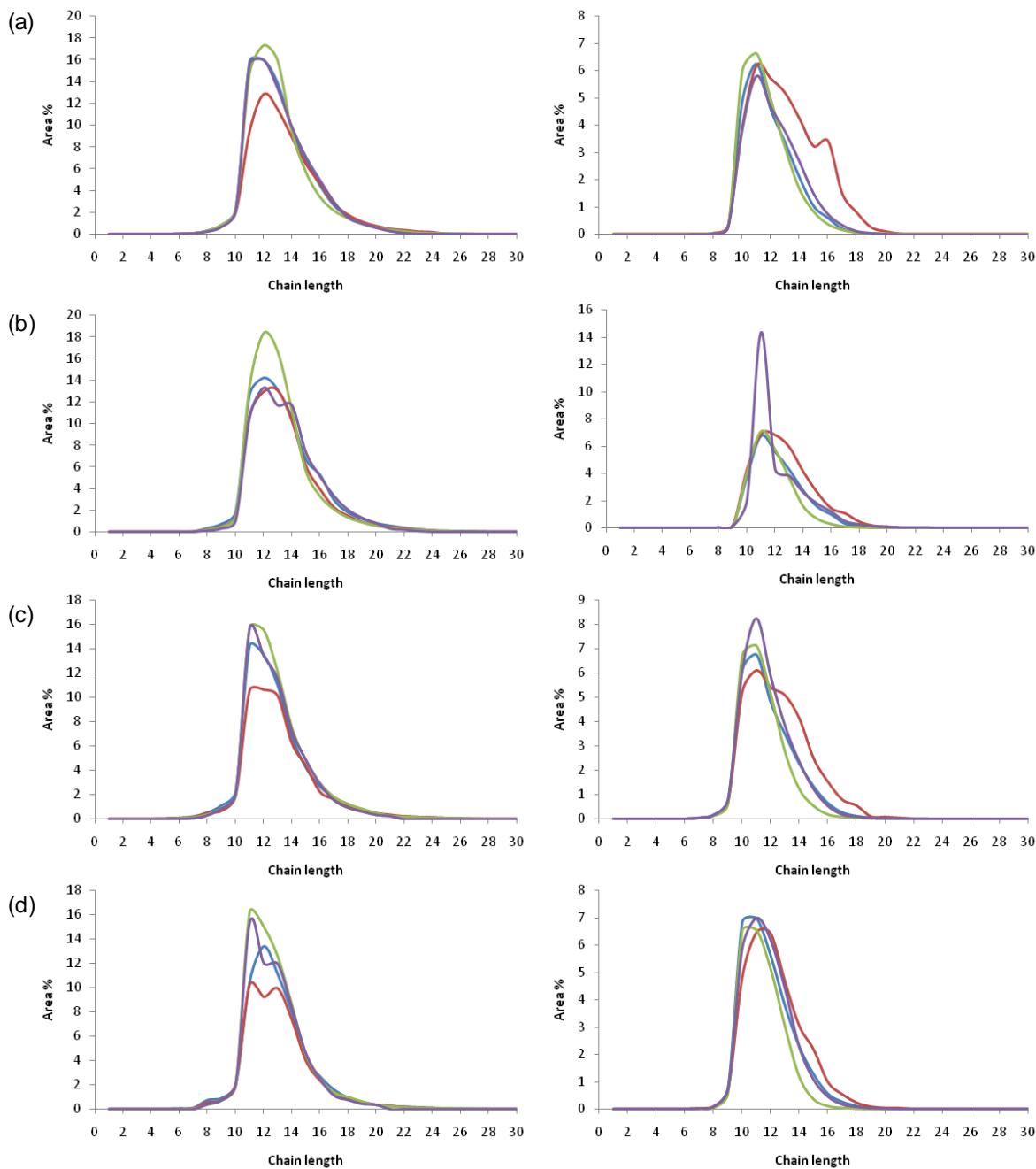


Figure 6.3 Comparison of the distribution of AIHC (left) and ArHC (right) in terms of chain length for samples III (a), IV (b), V (c) and VI (d). The results for column set 1 (—), 2 (—), 3 (—) and 4 (—) are compared.

The distribution of AIHC as a function of chain length was similar for all the column sets with only slight variations. The distribution of ArHC in terms of chain length was

similar for column sets 1, 3 and 4. It seems that column set 2 was more sensitive towards heavier ArHC. Despite the large amount of column bleed (observed in the same area of the chromatogram as these compounds), the reason for the increased sensitivity is not clear and further investigation is needed.

It was suspected that aromatic and heteroatomic compounds may contribute more towards improving lubricity compared to AIHC (see Chapter 2 and references therein). It is thus crucial to detect trace amounts of these compounds present in a matrix. Thus the information lost due to peak tailing in column sets 1, 3 and 4 might be critical for the purpose of this research. ArHC might also play a role in improving lubricity and thus the effective detection of these compounds are essential.

The separations obtained for each of the column sets were also compared in terms of the distribution of compounds in the two-dimensional separation space. Subsequently the *orthogonality* (and thus the amount of separation space utilised) for each of the column sets were compared mathematically.

2.1 *Graphical comparison*

The different column sets produced different distributions of classes of compounds in the plane of the two-dimensional chromatogram. The results obtained for each column set are summarised in Figure 6.4, 6.5, 6.6 and 6.7. In these figures, the location of AIHC is indicated with a pink outline, whilst the yellow outline indicates where the ArHC and Ox are located in the separation space. The ArHC and Ox are situated in the same region of the two-dimensional chromatogram for all the column sets. This is due to the fact that both types of compounds are relatively polar (or polarisable). Oxygenates interact with polar groups in the stationary phase through hydrogen bonding and dipole effects. Aromatic compounds tend to interact with polar groups in the stationary phase through π -interactions, i.e. induced dipole effects (see Chapter 3, Section 2.3). The column bleed is indicated with a black outline.

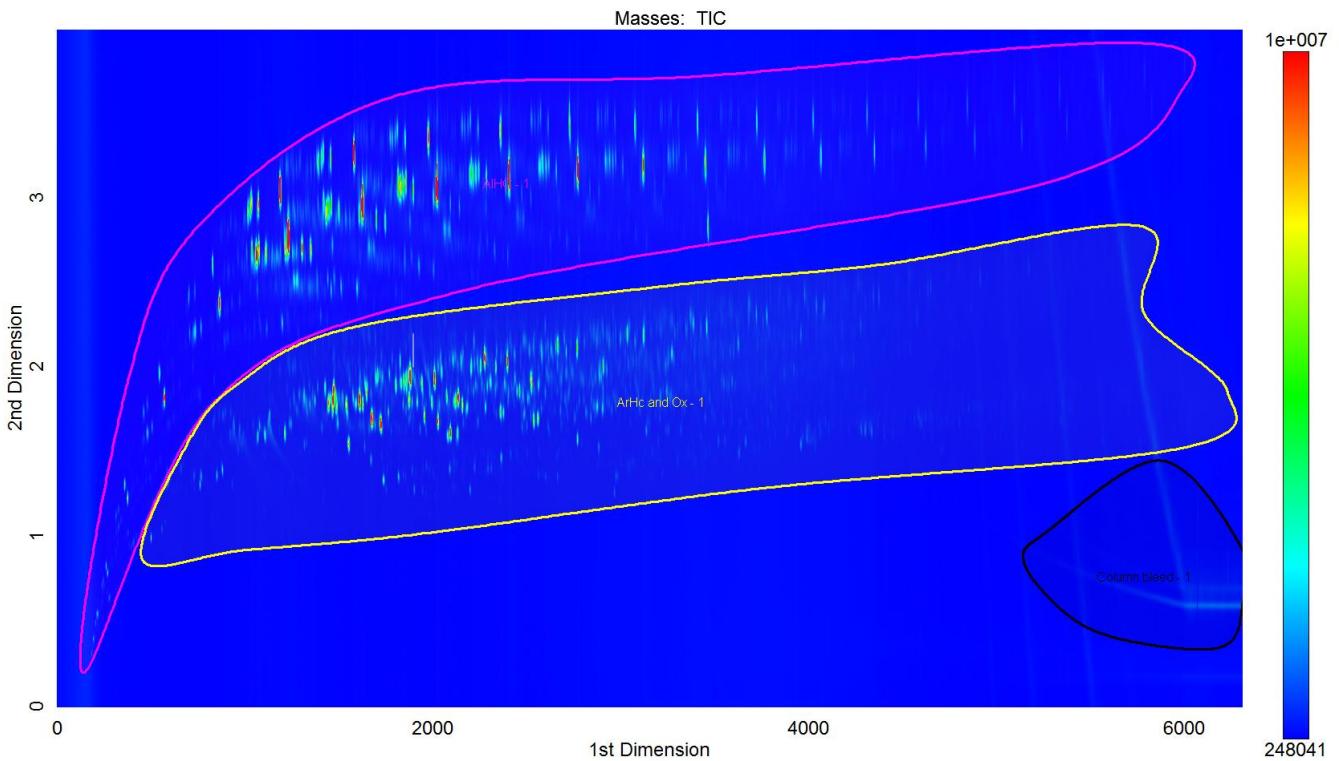


Figure 6.4 Distribution of classes of compounds (in terms of categories) for column set 1. The categories shown are AIHC (—) as well as ArHC and Ox (—) which elute in the same region. The column bleed (—) is also indicated.

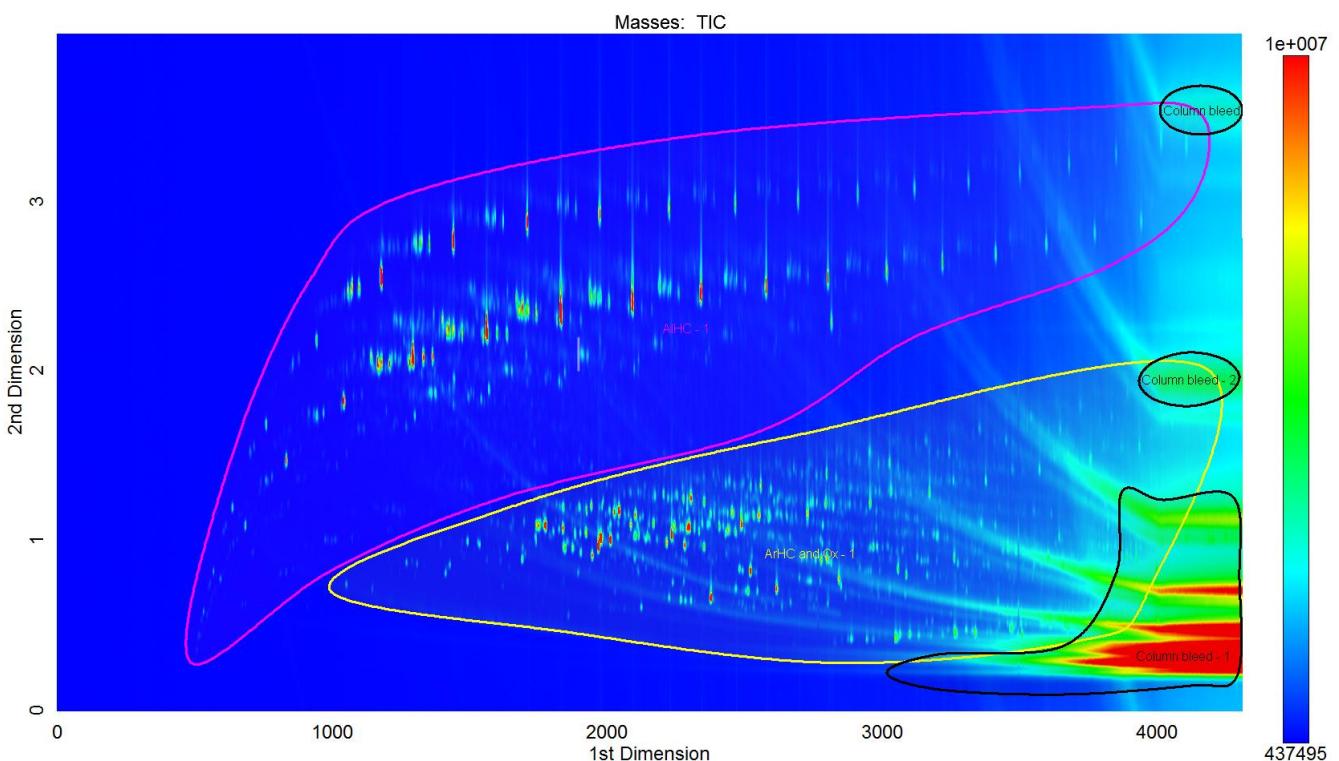


Figure 6.5 Distribution of classes of compounds (in terms of categories) for column set 2. The categories shown are AIHC (—) as well as ArHC and Ox (—) which elute in the same region. The column bleed (—) is also indicated.

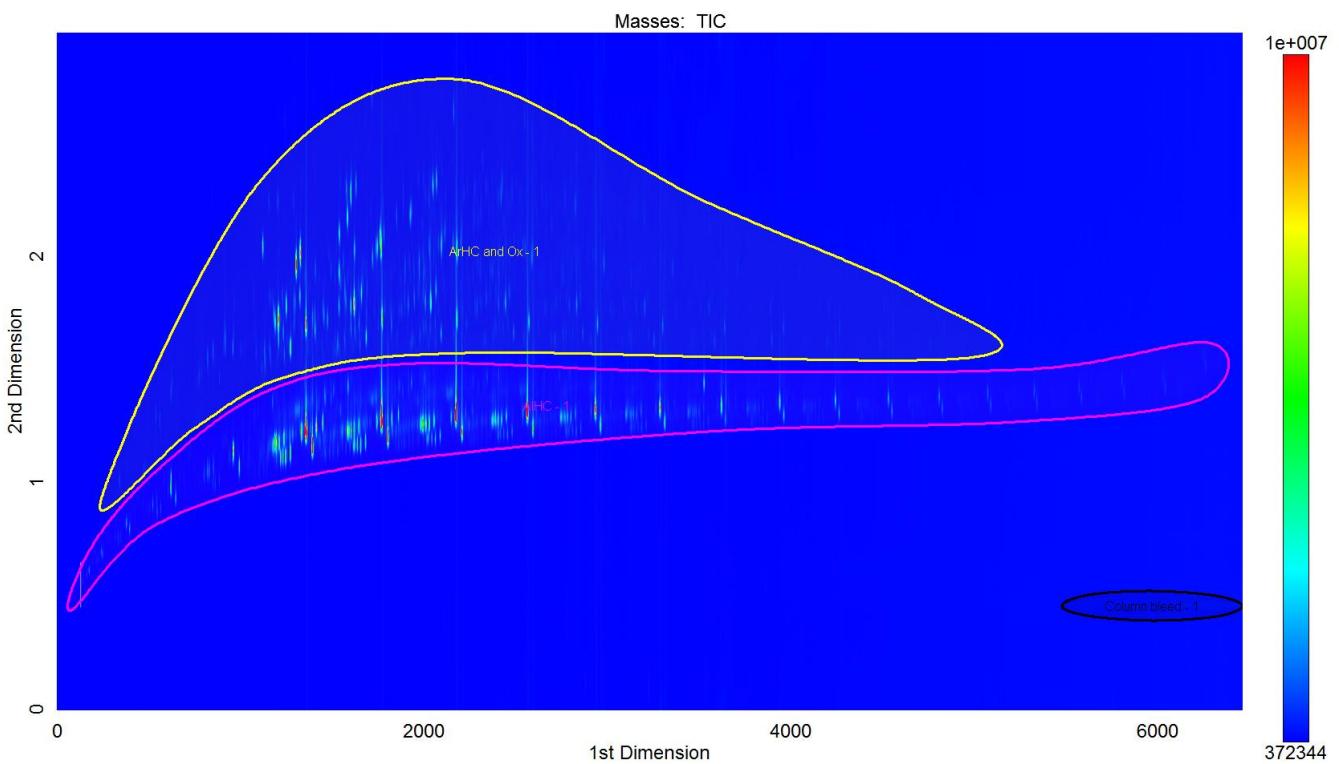


Figure 6.6 Distribution of classes of compounds (in terms of categories) for column set 3. The categories shown are AIHC (—) as well as ArHC and Ox (—) which elute in the same region. The column bleed (—) is also indicated.

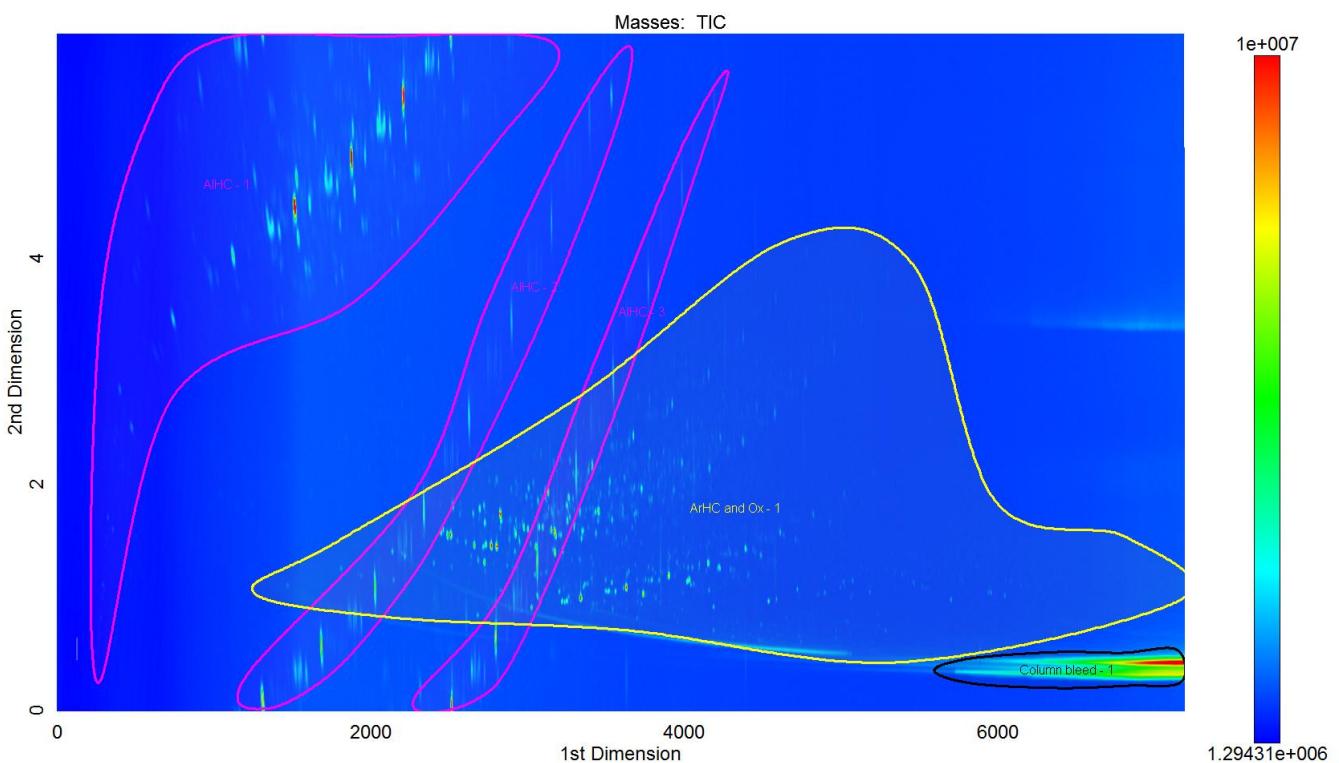


Figure 6.7 Distribution of classes of compounds (in terms of categories) for column set 4. The categories shown are AIHC (—) as well as ArHC and Ox (—) which elute in the same region. The column bleed (—) is also indicated.

For the “normal-phase” column set (Figure 6.6) the AIHC eluted earlier in the ^2D than the ArHC and Ox classes. The ArHC and Ox classes elute earlier in the ^1D . For the reversed column sets (Figure 6.4, 6.5 and 6.7) the opposite was observed, where more polar compounds elute later in the ^1D and earlier in the ^2D . Better class separation was observed for the reversed polarity column sets, as expected according to theory (see Chapter 5, Section 4.2).

The “reversed-phase” column sets can be compared in terms of degree of polarity in the ^1D (refer to Figure 5.2 for detailed chemical structures) since the ^2D was identical for these sets (see Table 5.2). The polarity of the ^1D of the column sets increase in the order $1 < 2 < 4$. As expected (see Chapter 5, Section 4.2) the distribution of compounds in the two-dimensional space increases as the polarity of the ^1D increases.

For column set 4, the AIHC are distributed throughout the ^2D to such an extent that severe wrap around occurs, and the heavier AIHC are wrapped around more than once (see red arrows in Figure 6.8). This could not be avoided, but was tolerated since these wrapped-around compounds could easily be distinguished from compounds that did not undergo wrap-around by their ^2D widths and mass spectra.

The ^1D of this column set is highly polar, and as a result the AIHC are not retained significantly (compared to the other column sets). These compounds are injected into the second column early on in the separation, when the secondary oven temperature is not very high. Since the ^2D is not very polar, these compounds are retained drastically at low temperatures and will take longer than the modulation period to elute from the column, as a result wrap-around occurs. To eliminate this wrap-around without the loss of resolution (of especially polar compounds), a slower rate can be used to increase the ^1D temperature. This will however result in an inconveniently long analysis time (see discussions in Chapter 5, Section 4.4.1).

The ^2D separation occurs essentially under isothermal conditions, and the general elution problem is encountered where the highly retained compounds undergo significant peak broadening. This peak broadening was used to distinguish between the peaks that were wrapped around and other peaks eluting in the same area. To

avoid this peak broadening (and consequently limit wrap-around without increasing analysis time), temperature programming in the second dimension would be very useful. This fast temperature programming is however not available at present.

It is interesting to note that the ionic liquid S used in the ^1D of column set 4, allows for the separation of heavy alkanes at low temperatures. Normally it is suggested to use a stationary phase with chemical characteristics similar to that of the target analyte (see Chapter 5, Section 4.2). Thus if heavy alkanes are to be investigated, a suitable stationary phase would be PDMS with only 5% or 1% of the methyl groups substituted by phenyl groups. For this reason the results obtained were compared with column set 3, which had a 5% phenyl PDMS column in the ^1D (Figure 6.8).

The results indicate that alkanes are less retained (but still sufficiently separated) in the ^1D of column set 4 compared to column set 3. The maximum temperature of the ^1D of column set 4 was 220 °C, whilst the ^1D of column set 3 was heated to 250 °C. For column set 4, the n-alkane with formula $\text{C}_{29}\text{H}_{60}$ eluted when the ^1D temperature was around only 160 °C. For column set 3 the largest alkane observed was $\text{C}_{27}\text{H}_{56}$.

The most challenging aspect of working with column set 2 was the large amount of column bleed observed. Many of the polar compounds elute in the same area as the column bleed and thus classification of these compounds is complicated. Nevertheless better class separation was observed compared to column sets 1 and 3, where almost no column bleed was observed. Column set 4 is superior to column set 2 in terms of column bleed, and no separation is lost since the ^1D is highly polar. In fact the separation achieved with column set 4 might be better than that of column set 2, but the wrap-around discussed above is not ideal. The temperature difference between the primary and secondary ovens was kept high throughout the analysis, in an attempt to limit wrap-around, but because of this the separation of ArHC and Ox was not optimum and the latter class resolution was not achieved.

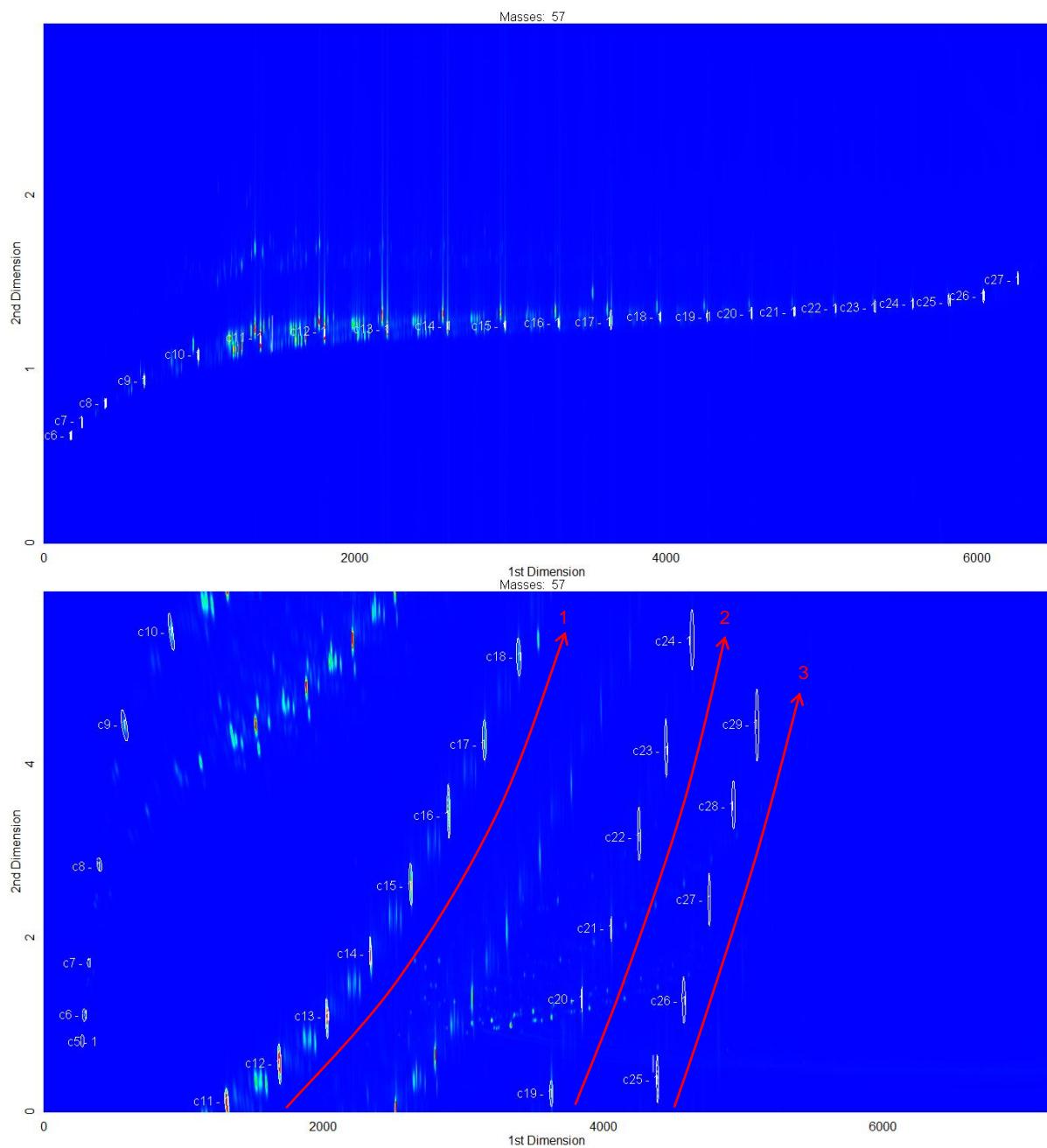


Figure 6.8 EIC of mass 57, indicating the n-alkanes observed in column set 3 (top) and column set 4 (bottom) for the same sample. The red arrows and corresponding numbers indicate how many times wrap-around occurred.

The separation of certain compounds was difficult due to their chemical similarity. As mentioned before, alkenes and cyclic alkanes elute in the same areas of the two-dimensional separation space and have very similar mass spectra. Thus even with the use of a mass selective detector, it is difficult to distinguish between these compounds. The same argument holds for compounds that have higher degrees of unsaturation, e.g. alkynes, dienes, cyclic alkenes and bicyclic alkanes. It was observed that these compounds could be separated to a certain degree depending

on the polarity of the ^1D . This is illustrated in Figure 6.9 where extracted ion chromatograms (EIC) were used to see all the isomers of $\text{C}_{11}\text{H}_{22}$ ($\text{HDI} = 1$, thus containing one double bond or cyclic structure) present in a sample for the different column sets. Figure 6.10 shows the EIC for the isomers of $\text{C}_{11}\text{H}_{20}$ ($\text{HDI} = 2$, thus containing a triple bond, two double bonds, two cyclic structures or a double bond together with a cyclic structure).

The unsaturated compounds are separated from compounds containing saturated rings, because the more polarisable electrons in the π -bond can interact more readily with the polar stationary phase. As a result the unsaturated compounds will elute later in the ^1D and (slightly) earlier in the ^2D for the reversed column sets.

The degree of separation between the subclasses: cyclic alkanes and alkenes, increases as the ^1D polarity increases. Sufficient separation is observed for both column sets 2 and 4 although better separation is observed for column set 4. The $\text{C}_{11}\text{H}_{20}$ isomers were only separated into discernable rooftiles (groups) where using column set 4. This is probably due to the fact that an HDI of two gives rise to many more possible isomers compared to an HDI of one. Consequently a more polar ^1D column is necessary to resolve these subclasses. This was also investigated mathematically and will be discussed in the following section.

To obtain this type of separation usually requires an LC pre-separation step followed by GCxGC analysis [2]. By using an ionic liquid as S in the ^1D of GCxGC no pre-separation steps are required to separate cyclic structures from unsaturated structures. The separation of cyclic olefins (alkenes) and dienes (both with $\text{HDI} = 2$) have been reported using a column set similar to column set 2 (GCxGC with PEG in ^1D) [3]. Note that a long modulation period of ten seconds was used to achieve this separation.

Up to now the separation of compounds (and classes of compounds) in the two-dimensional separation space were discussed based on generally observed features. These differences in separation ability of each column set can be quantified in terms of orthogonality (see Chapter 3, Section 4.2) as discussed in the following section.

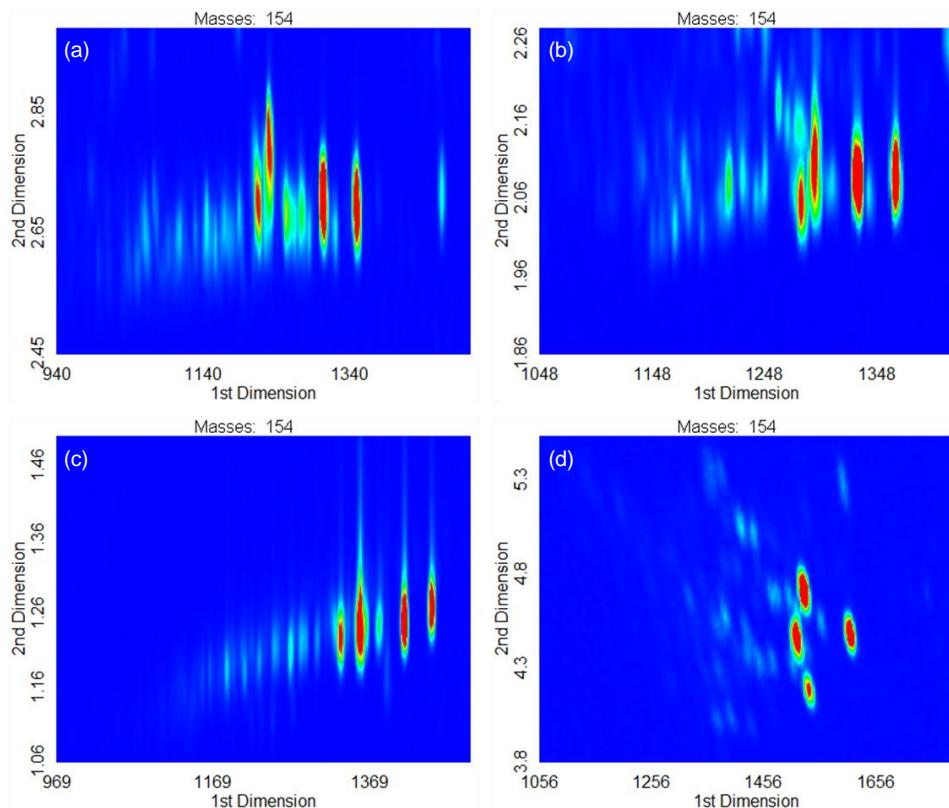


Figure 6.9 EIC of mass 154 for the comparison of the separation of $C_{11}H_{22}$ isomers. The separation achieved for the same sample in column set 1 (a), 2 (b), 3 (c) and 4 (d) are shown. Rooftiles seem better defined for (b) and especially (d).

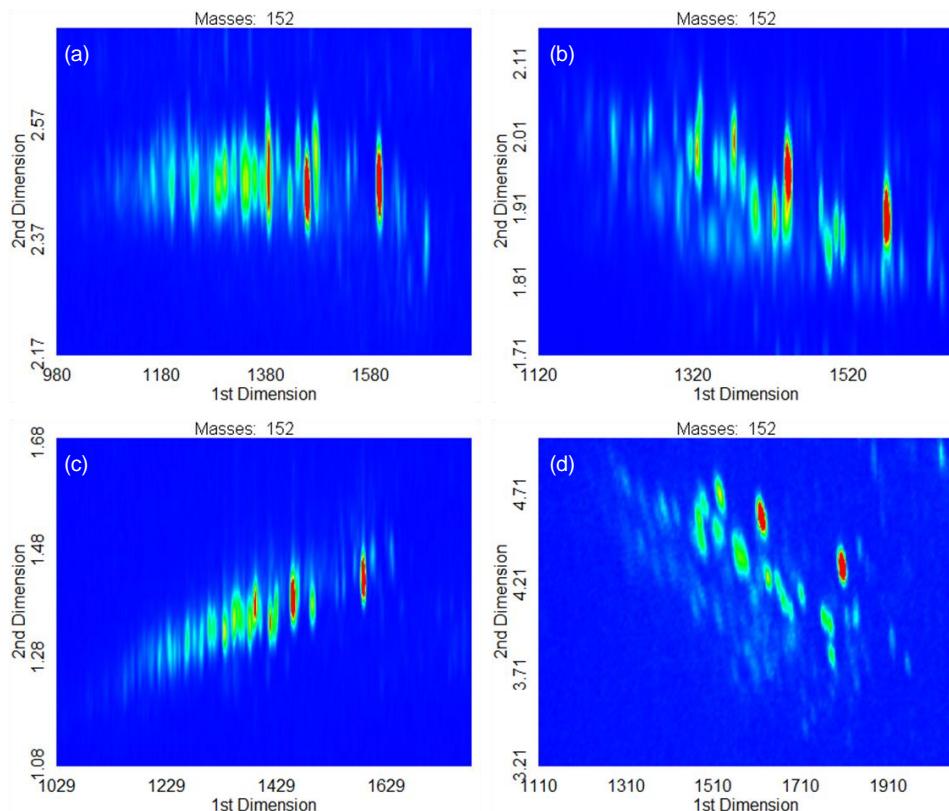


Figure 6.10 EIC of mass 152 for the comparison of the separation of $C_{11}H_{20}$ isomers. The separation achieved for the same sample in column set 1 (a), 2 (b), 3 (c) and 4 (d) are shown. Rooftiles seem better defined for (b) and especially (d).

2.2 Mathematical comparison

The separation and detection of ArHC and especially Ox was crucial for the purpose of this research, since there was no indication that AIHC will contribute to lubricity. For this reason, the separation abilities of the various column sets were investigated in terms of only the ArHC and Ox.

The distribution of compounds in the two-dimensional space is directly related to the *orthogonality* of the system. The *orthogonality* of a system is the inverse of the linear association present between 1t_R and 2t_R (see Chapter 3, Section 4.2). Since the compositions of the samples vary, the separation space utilised by each sample will also vary. This means that orthogonality does not only depend on the columns used in the 1D and 2D , but also on the sample analysed. Table 6.2 shows the correlation coefficients determined for the retention times of ArHC and Ox, whilst ignoring the presence of AIHC. A confidence limit of 95 % was used.

Table 6.2 The correlation coefficients indicating the linear association between 1t_R and 2t_R . These values are inversely related to the *orthogonality* of the system in terms of each specific sample. The smallest correlation, indicates which column set is the most orthogonal.

Sample	Column set 1	Column set 2	Column set 3	Column set 4	Smallest correlation
I	0.74	0.72	0.95	-0.12	4
II	0.73	0.41	0.97	0.25	4
III	0.26	0.09	0.34	-0.17	4
IV	0.29	0.14	0.20	-0.08	4
V	0.25	0.03	0.37	-0.21	4
VI	0.25	-0.05	0.36	-0.18	2
VII	0.72	0.23	0.83	0.34	2
VIII	0.75	0.40	0.80	-0.01	4
IX	0.37	-0.01	0.73	-0.19	2

The easiest way to understand correlation is in terms of an ellipse, where its width indicates the degree of association along its length (see Chapter 4, Section 5). The direction (or slope) of the length of the ellipse indicates whether the association is positive or negative. When a system is not orthogonal, compounds will be arranged

on the diagonal of the separation space as explained in Chapter 3, Section 4.2. When using correlation coefficients to compare the orthogonality of several systems, the aim is to measure the width of an ellipse along this diagonal line. This diagonal exists in two-dimensional GC systems, because there is always an underlying dependence on the boiling points of the analytes in both dimensions.

In cases where a negative correlation coefficient was obtained, it means there was a negative association, i.e. the ellipse along the diagonal described above became so wide, that its length is mathematically seen as its width and the width along the diagonal is now its length. The correlation coefficient no longer describes the variation along the diagonal. This might indicate that the underlying dependence on the boiling point becomes less important and almost negligible for the 1D column separation, although further investigation is needed.

Negative values will typically be observed in samples where there are a few classes (rooftiles) containing a lot of isomers, which are well separated from one another. A positive value for such a system might be obtained when a wider variety of compounds classes are present.

The results indicate that column set 2 and 4 represents the most orthogonal conditions for the separation of ArHC and Ox, when AIHC is excluded. For samples VI and IX, negative correlation coefficients were obtained for both column sets 2 and 4. To compare and interpret the separations achieved for these samples, the scatterplots (of 1t_R vs. 2t_R) were compared visually in Figure 6.11.

For column set 2, better class separation was obtained (rooftiles are more readily distinguishable), whilst for column set 4, better separation of categories and subcategories were achieved (e.g. for sample IV the three main groups of peaks are separated to a greater extent from one another compared to the separation achieved for these groups on column set 2).

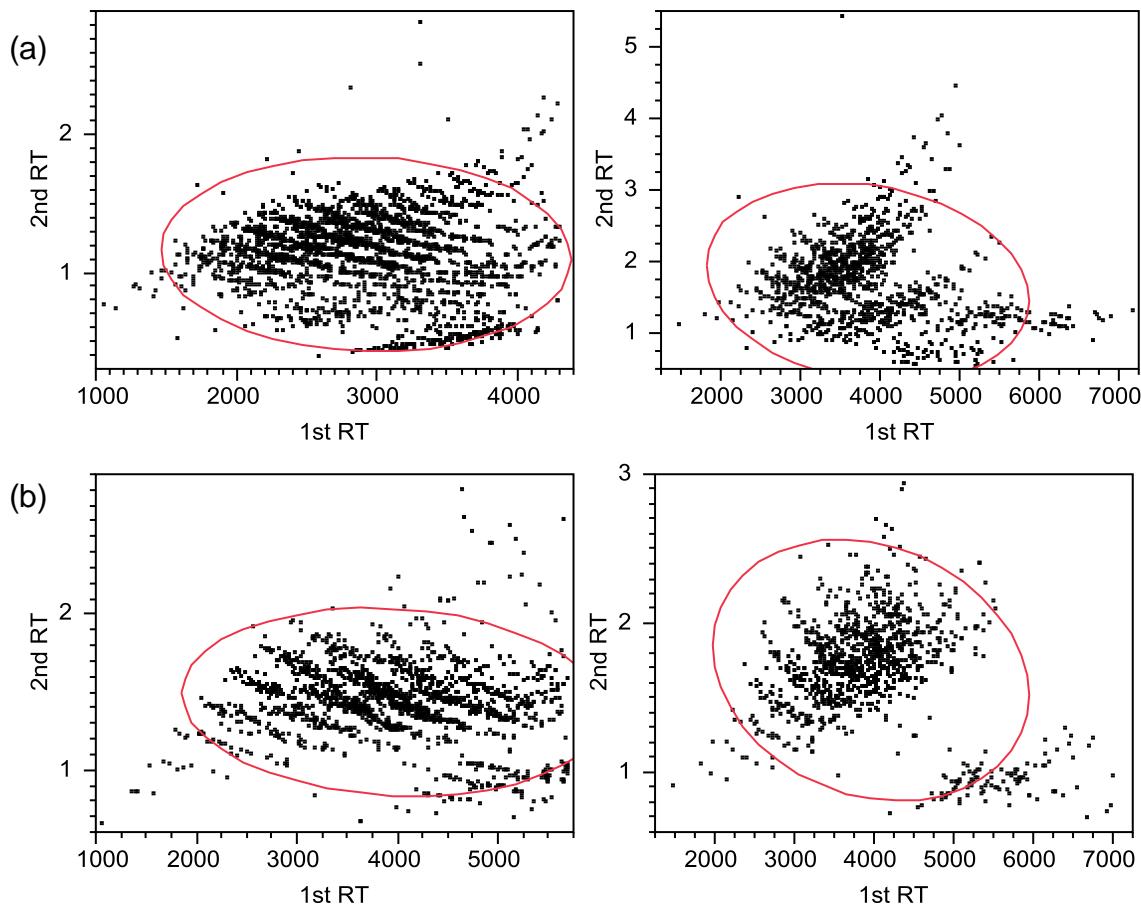


Figure 6.11 Comparison of the distribution of peak apices (of only ArHC and Ox species) along the diagonal of the two-dimensional separation space for column sets 2 (left) and 4 (right). Results for samples VI (a) and IX (b) are shown.

The separation of cyclic and unsaturated compounds (for specifically $C_{11}H_{22}$ and $C_{11}H_{20}$) was also investigated using a similar mathematical approach as discussed above. Mass spectra of peaks within the specified region of the chromatogram corresponding to Figures 6.9 and 6.10 respectively were compared to mass spectral libraries. A minimum similarity of 700 out of a possible 1000 was required to tentatively assign the identity of a peak. The compounds with masses 154 (for $C_{11}H_{22}$) and 152 (for $C_{11}H_{20}$) were used in further calculations. Table 6.3 shows the correlation coefficients determined for the retention times of these compounds. A confidence limit of 95 % was used.

The negative values are only indicative of a negative slope (compare figures 6.9 and 6.10) in this instance and therefore values closer to "0" represent a better separation. The results indicate that isomers of $C_{11}H_{22}$ and $C_{11}H_{20}$ are separated better on column sets 2 and 4 compared to column sets 1 and 3, since a larger area of the

defined separation space is occupied. Taking the correlation coefficients into account together with the observations made after visual inspection (see Chapter 6, Section 2.1), it can be deduced that it is beneficial to use a more polar stationary phase in the ^1D of GCxGC for the separation of cyclic and unsaturated isomers. The ionic liquid column used in column set 4 showed a lot of promise in this regard.

Table 6.3 The correlation coefficients indicating the linear association between $^1\text{t}_R$ and $^2\text{t}_R$ of compounds with masses 154 and 152. These values are inversely related to the *orthogonality* and also the degree of separation achieved. A smaller value indicates that a larger area of the two-dimensional separation space is utilized.

Column set	Mass 154 ($\text{C}_{11}\text{H}_{22}$)	Mass 152 ($\text{C}_{11}\text{H}_{20}$)
1	0.46	0.48
2	0.13	-0.067
3	0.69	0.70
4	0.073	-0.18

3. Conclusion

The only question that remains is: Which column combination is the best for the purposes of this research? It is very important to keep the purpose of the research in mind when selecting a column set. As seen from above discussions, the different column sets all have advantages and disadvantages. Column set 4 represents unique characteristics, and better compound resolution is obtained, but at the expense of class resolution. The severe wrap-around of AIHC was tolerated, but this column set is less ideal for the general analysis of petrochemical samples. It can be concluded that column set 2 has the most advantages in terms of what needs to be achieved, despite the large amount of column bleed that is observed. This includes efficient separation and detection of Ox and ArHC, since these compounds are deemed important for lubricity. Henceforth the results obtained from column set 2 will be used.

References

1. Armstrong, D.W.; Payagala, T.; Sidisky, L.M. *LC-GC Europe* **2009**, 22, pp.459 – 467.
2. Van der Westhuizen, R.; Potgieter, H.; Prinsloo, N.; De Villiers, A.; Sandra, P. *Journal of Chromatography A* **2010**, 1218, pp. 3173 – 3179.
3. Van der Westhuizen, R.; Crouch, R.; Sandra, P. *Journal of Separation Science* **2008**, 31, pp. 3423 – 3428.

Chapter 7: Statistical Analysis of Results

Contents

1.	A look at the lubricity.....	112
2.	Mixture models	113
3.	Linear associations	116
4.	Discussion	118

1. A look at the lubricity

To get an idea of how these results can be related to the lubricity of the diesel samples, a ternary diagram was used to indicate the different relative compositions of the samples and the markers were coloured on a continuous scale according to the WSD values (see Figure 7.1). The lubricity was measured in terms of WSD (see Chapter 2, Section 3.5) and a smaller WSD indicates better lubricity. To simplify interpretations, the composition of the samples were summarised in terms of the total amounts of AIHC, ArHC and Ox present (see Chapter 6, Table 6.1).

The ternary diagram indicates that the samples containing more Ox have smaller WSD values and thus better lubricity. A larger ArHC content also improves lubricity, but to a lesser extent. In order to investigate the relationship between relative composition and WSD mathematically, a mixture model analysis was performed.

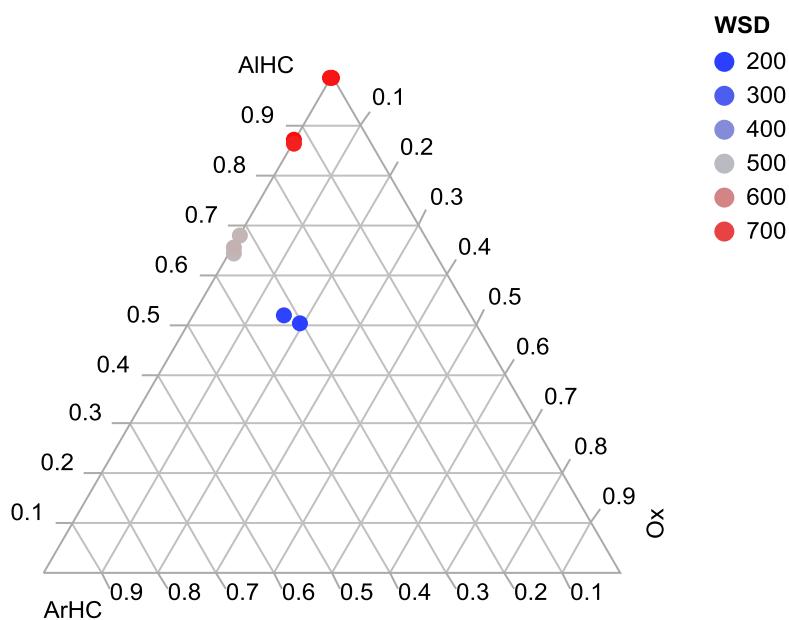


Figure 7.1 A ternary diagram indicating the relationship between the composition and WSD using results obtained from column set 2.

2. Mixture models

The relationship between lubricity and chemical composition was statistically investigated using a mixture model as explained in Chapter 4, Section 6. A prediction equation (Equation 7.1) was obtained using JMP® 8.0.1, by defining the variables as mixture components and using a least squares approach. The response surface was a saddle point. Due to the limited number of diesel streams available, this model was seen as provisional and further investigation is needed to find a more accurate prediction equation. Even though this is not ideal and no definite conclusions can be made, the results give an indication of which components contribute to lubricity.

$$\begin{aligned} \text{WSD} = & 777 \text{ (AIHC)} - 2290 \text{ (ArHC)} - 16567 \text{ (Ox)} + 3419 \text{ (AIHC)(ArHC)} \\ & + 70740 \text{ (AIHC)(Ox)} - 62190 \text{ (ArHC)(Ox)} \end{aligned} \quad (7.1)$$

The coefficients on the ArHC and Ox are both negative indicating that the WSD decreases as the relative amounts of these categories increase and thus these categories might improve lubricity. The slope of the response surface is much larger along the Ox axis compared to the ArHC axis, indicating that Ox might play a bigger role in improving lubricity. The interaction between ArHC and Ox (last term in Equation 7.1) also has a negative relationship with WSD (negative coefficient) indicating that as the relative amount of these categories increase, the WSD decreases and lubricity might be improved.

The model was fitted against the ternary diagram showing the compositions of the different samples shown in Figure 7.2. The grey area indicates the constraints in the data, i.e. the combinations of components for which a response (WSD) was not available. The red line indicates the contour line for a WSD of 460 µm (which is the maximum allowed value for fuels to be sold commercially) based on the response surface described by Equation 7.1. The red dots indicate the direction in which the WSD will increase according to the model. This means that samples with compositions below this line should have WSD less than 460 µm and consequently sufficient lubricity.

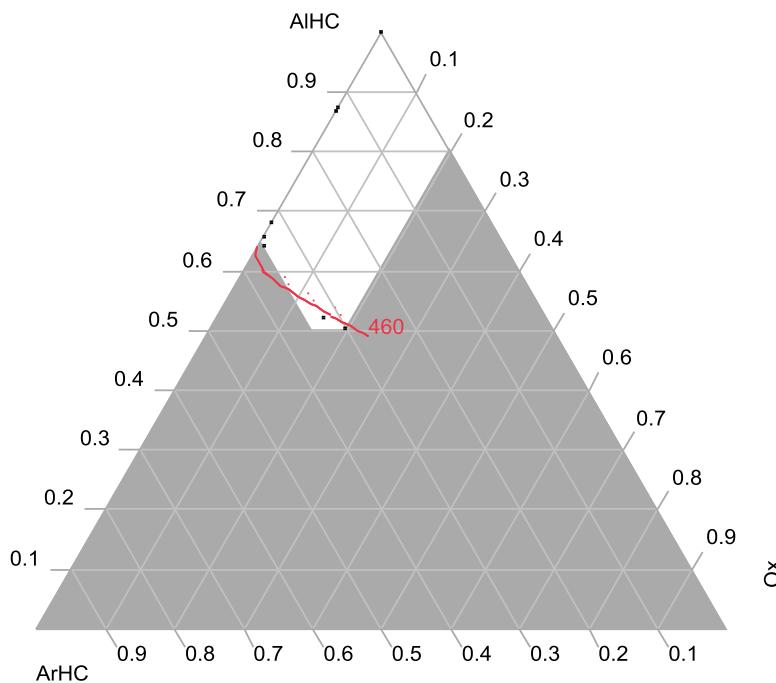


Figure 7.2 A mixture model indicating the relationship between composition and WSD. The black dots represent the compositions of the available samples. The grey area indicates the constraints on the sample compositions. The red line is the contour line at WSD = 460 μm . The WSD increases in the direction of the red dots.

It was also necessary to consider the effect of chain length. It is suspected that compounds with less than 12 carbon atoms will have a smaller influence on lubricity than larger compounds. The choice of this specific cut-off point was based on compounds used as lubricity additives (see Chapter 2, Section 3.3 and Chapter 5, Section 5). The three categories (AIHC, ArHC and Ox) were divided into two groups each according to chain length (indicated by <12 or ≥ 12). These values are reported in Table 7.1.

To represent this data graphically and to compare the effect of chain length, ternary diagram were used where the smaller (<12 C's) and larger (≥ 12 C's) compounds in a category (AIHC, ArHC or Ox) were compared to the total amount of all the other components (indicated as "Other"). The results are given in Figure 7.3. It was not possible to fit a model to the data, but the markers were again coloured according to their respective WSD values in order to make interpretations possible.

Table 7.1 Composition of samples in terms of chain length.

Sample	AIHC		ArHC		Ox	
	<12	>=12	<12	>=12	<12	>=12
I	49.91	49.81	0.26	0.02	0.00	0.00
II	46.34	53.46	0.17	0.03	0.00	0.00
III	12.45	51.71	10.47	24.45	0.82	0.10
IV	13.00	52.53	11.48	22.79	0.18	0.03
V	13.82	38.14	12.21	20.18	12.70	2.96
VI	13.63	36.46	12.07	18.11	14.99	4.45
VII	84.74	1.88	13.25	0.13	0.00	0.00
VIII	86.66	0.59	12.75	0.00	0.00	0.00
IX	31.50	36.59	13.54	18.37	0.00	0.00

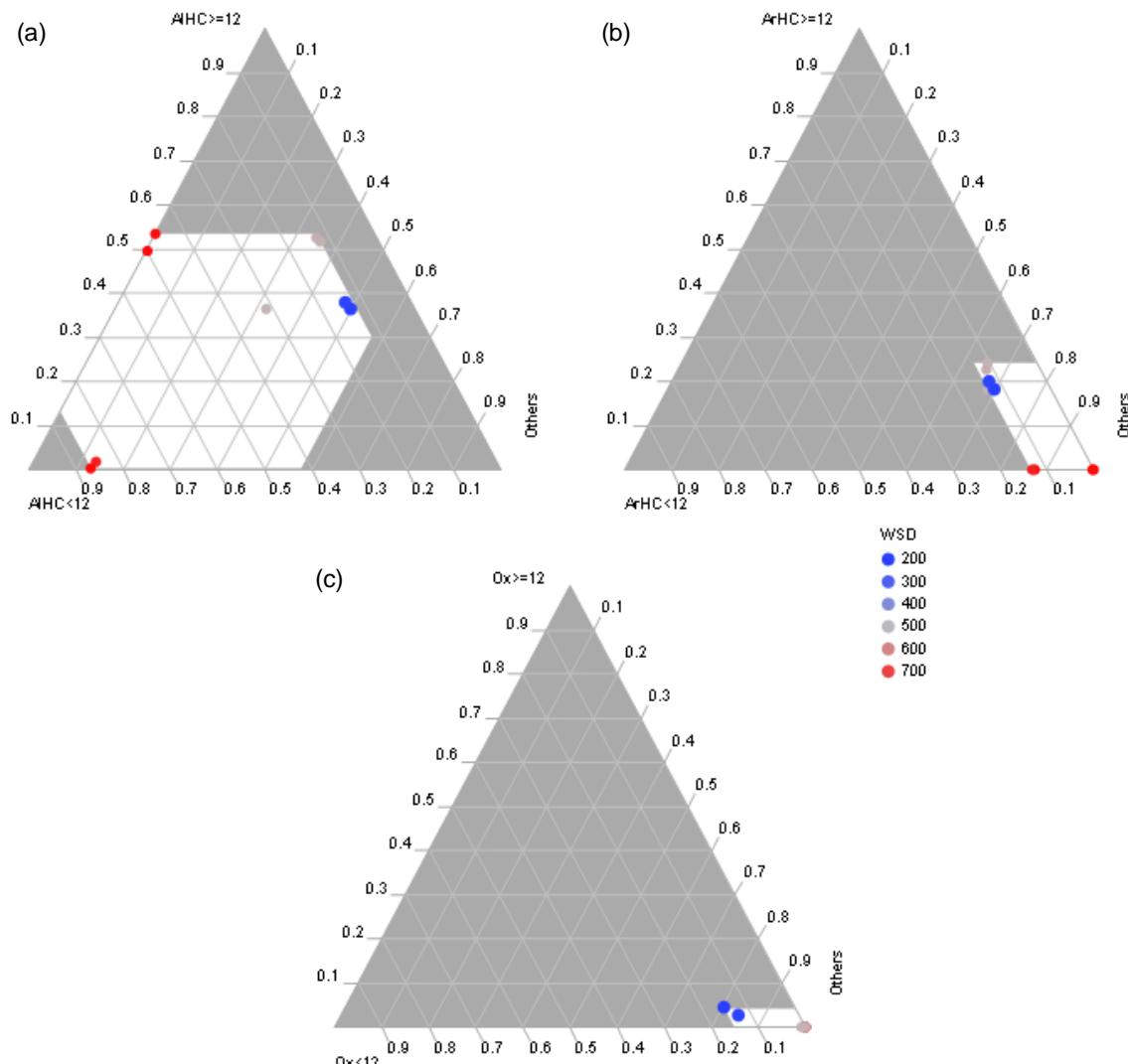


Figure 7.3 Influence of chain length on WSD. The differences between shorter and longer compounds were compared for each of the three categories; ArHC (a), AIHC (b) and Ox (c).

The results indicate that WSD increases as AIHC<12 and AIHC>=12 increases. The WSD decreases as ArHC>=12 increases, whilst a clear trend is not observed for ArHC<12. This means that the presence of larger ArHC could improve lubricity. An increase in both Ox<12 and Ox>=12 results in a decrease in WSD. A smaller increase in Ox>=12 is needed to decrease the WSD by a certain amount compared to Ox<12, indicating that Ox>=12 will improve lubricity to a greater extent, even though Ox<12 could also improve lubricity.

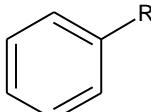
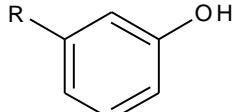
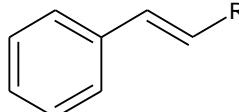
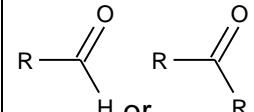
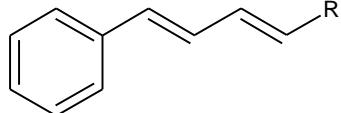
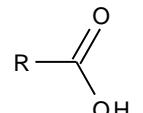
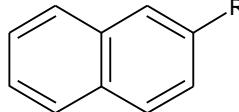
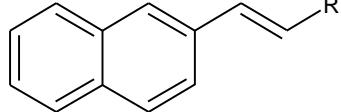
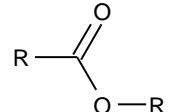
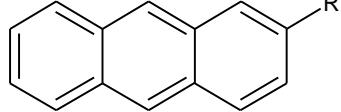
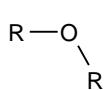
The information gained from these results are valuable, but more detailed information may be necessary, e.g. which specific Ox species will improve lubricity. It is possible to predict a model based on a mixture consisting of more than three components, but the results are not as easy to interpret graphically.

3. Linear associations

The linear association between the compound classes and WSD was also investigated in an attempt to gain more specific information. From above results, it can be assumed that AIHC will not improve lubricity and thus these compound classes were not investigated further. The correlation coefficients are summarised in Table 7.2 below.

The results show that all the compound classes are negatively associated with WSD. This means that these compound classes may improve lubricity. Even though the results indicate a strong linear association between the compound classes and WSD, this does not necessarily mean that these compound classes improve lubricity. Experimental proof is needed to conclude that a causal relationship exists (see Chapter 4, Section 5.1).

Table 7.2 Correlation coefficients (r) of ArHC and Ox with WSD. The R-group represents alkyl groups, and the non-aromatic double bonds rings represent the degree of unsaturation present in the molecule, thus it can be saturated ring structures as well.

ArHC			Ox		
General formula	Example	r	Functional group	Example	r
C_nH_{2n-6}		-0.46	phenol		-0.87
C_nH_{2n-8}		-0.76	aldehyde / ketone		-0.86
C_nH_{2n-10}		-0.96	carboxylic acid		-0.85
C_nH_{2n-12}		-0.92	alcohol	$R-OH$	-0.85
C_nH_{2n-14}		-0.49	ester		-0.80
C_nH_{2n-18}		-0.85	ether		-0.71

A larger $|r|$ value indicates a stronger linear association. When considering the Ox compound classes, it can be observed that compounds containing phenol, aldehyde, ketone, carboxylic acid or alcohol functionalities have the strongest linear association with WSD. For the ArHC, compounds with the general formulas C_nH_{2n-10} and C_nH_{2n-12} (with an overall HDI = 6), have the strongest linear association with WSD. Compounds containing three aromatic rings, such as anthracene, are also strongly associated with WSD.

4. Discussion

Even though the available data set is insufficient for a detailed statistical analysis, some conclusions could be made regarding the effect of individual components on lubricity. The results indicate that the presence of Ox and ArHC in a mixture can improve its lubricity whilst AlHC does not contribute significantly. This corresponds well with literature (see Chapter 2, Section 3 and references therein) which indicates that compounds that can attach to a metal surface (through chemisorption or physisorption) will prevent wear by forming friction reducing layers during boundary lubrication.

It was also shown that larger Ox (containing 12 or more carbon atoms) seem to have a more significant influence on lubricity than smaller Ox. This is supported by the hypothesis that long carbon chains assemble like bristles whilst the polar regions of these compounds adsorb onto the metal surface (see Chapter 2, Section 3.1.1). Further experimental proof is needed to deduce causality.

The investigation of the linear association between WSD and certain compound classes indicated that specific compounds are more strongly associated with lubricity. It might be possible that these classes are highly associated with another underlying characteristic of the sample (or an undetected compound) which causes better lubrication. Experimental proof is thus needed to prove whether a linear relationship exists between these compound classes and lubricity.

Chapter 8: Conclusions and final remarks

Contents

1. Main conclusions	120
2. Additional Conclusions	122
3. Recommendations.....	122

1. Main conclusions

GCxGC-TOFMS proved to be an invaluable tool for the analysis of diesel samples. The enhanced peak capacity allowed for these complex samples to be separated into their individual components. The structured chromatograms made a group-type analysis possible which was sufficient for the purposes of this project and much simpler than identifying individual compounds. The focus of this work was to separate and identify compounds present in diesel streams that may contribute to fuel lubricity. It was suspected that more polar compounds such as oxygenates and, to a lesser extent, aromatic hydrocarbons contribute to lubricity. For this reason separations were optimised and compared with the focus on separation of polar species.

A qualitative analysis was performed, since the relative amounts of components were sufficient for the purposes of the research. The compounds were assigned to classes manually, based on the 1t_R and 2t_R as well as mass spectral data. In some cases, peak shapes were also used for the identification of compounds. Even though no internal standards were analysed, the results obtained with the different column sets were similar. The separation methods will not be transferable to FID, because incomplete separation can only be addressed with the additional dimension of separation that mass spectrometry offers. Peak deconvolution is not possible when only FID data is available. Attempts to simultaneously determine the bulk composition of the samples together with trace compounds were not as effective.

Only compounds with a S/N of 500 or higher were considered due to the complex nature of the samples. There exists a possibility that some compounds at lower levels may also contribute to lubricity, and it is suggested that an alternative approach be followed to classify compounds into classes, e.g. by using algorithms based on mass spectral data. When assigning classes manually (the approach followed in this study), it is impractical to consider lower S/N ratios.

Four different column combinations were investigated and compared. In cases where the ^1D column was more polar than the ^2D column, better class separation was achieved (as expected). It was also shown that the degree of polarity in the ^1D influences how much of the two-dimensional separation space is utilised. Column sets with more polar columns in the ^1D , separated ArHC and Ox better, since these systems were more orthogonal compared to other column sets.

It was found that column set 2 (PEG stationary phase in ^1D) gave a more representative sample composition, since satisfactory class resolution was obtained and no peak tailing (of polar compounds) was observed in the ^1D , despite the large amounts of column bleed. The separation and detection of ArHc and Ox were the most important factors when considering which column combination to use, since it was suspected that these compounds could improve the lubricity properties of a fuel.

Based on lubricity measurements (HFRR), it can be deduced that some of the compounds contributing to lubricity are lost during final refining processes, since the feed streams to these processes have much smaller WSD values and the final blending streams have WSD values larger than 460 μm .

The substantial amount of information gained from the GCxGC-TOFMS analyses had to be reduced during statistical calculations due to the limited number of diesel streams. As a result it was not necessary to utilise alternative analysis methods such as NMR, IR and SFC as proposed initially. The amount of sample streams available (sample size, n) was insufficient for a thorough statistical analysis, but some trends were still identified based on a tentative model. The positive effect that both Ox and ArHC (individually and combined) have on lubricity was illuminated. The model also indicated that the amount of Ox play a more important role (compared to ArHC) towards improving lubricity. It was also noted that longer chain Ox possibly contribute more to lubricity than shorter chain Ox.

Some statistically significant linear associations were observed between lubricity and specific types of Ox and ArHC. It is necessary for these linear associations to be investigated experimentally (i.e. by noting changes in WSD with an addition to a

matrix of small amounts of these selected compounds), before it can be concluded that these specific compounds cause an increase in lubricity.

2. Additional Conclusions

The ionic liquid S used in the 1D of column set 4 has only recently become commercially available and has up to now primarily been used to analyse biodiesel samples. This column presented interesting separation characteristics for AIHC. It was observed that this column could be used for the analysis of heavy alkanes at low temperatures, since they are not retained significantly. The ability of this column to separate unsaturated compounds from compounds with saturated ring structures was also illustrated. This separation cannot be achieved as well with other columns or with the use of mass selective detectors. When such a polar column is used in the 1D of a GCxGC system, temperature programming, unavailable at present, might be useful in the second dimension.

Even though column set 4 presented unique separation of AIHC, the separation of ArHC and Ox was not ideal. This was because the improved compound separation resulted in the overlap of compound classes. Another drawback of this column set was the severe peak tailing observed for highly polar oxygenates, such as carboxylic acids, a feature mentioned in the column specifications.

3. Recommendations

For future work it is recommended that an experimental design be used to ensure sufficient data is available for a detailed statistical analysis. A design of experiments (DOE) approach is based on previously attained data and can give accurate descriptions of how many different fuel samples are necessary and how much their chemical compositions should vary. The results can also be used for a more empirical approach where “fuels” with different compositions can be made by mixing specified amounts of pure components, followed by detailed statistical analyses. By

doing this, it might be possible to get more information regarding what types of Ox and ArHC contribute to lubricity.

In order to detect polar species, it could be advantageous to pre-concentrate the polar fraction of the samples. By doing this, compounds that were at S/N levels lower than 500 in the fuel sample (as analysed by GCxGC) can be investigated (these compounds were not considered in the present study due to sample complexity) since the bulk of the sample (saturates) can be removed. To concentrate the polar fractions, extraction methods (e.g. using methanol or ionic liquids) or chromatographic methods (e.g. SFC) can be used.

In future focus should also be on lower levels of polar compounds present, since this research supports literature indications that polar compounds are important for lubricity. It is not necessary to spend time on quantifying and qualifying AIHC present in diesel streams. Performing further data mining on existing data sets should be considered. Setting up calibration curves based on certain extracted ions (characteristic of specifically oxygenates) and correlating these amounts with WSD values should present useful information.

It will also be interesting to determine the effect of certain identified compounds on lubricity by spiking a low lubricity fuel with small amounts of these compounds and monitoring whether a change in WSD occurs. By doing this causality can be proven and the minimum amount (of certain important compounds) needed to improve lubricity sufficiently can also be determined.

Appendix A: Two-dimensional chromatograms of samples

Column set 1:

Sample I.....	A-2
Sample II.....	A-2
Sample III.....	A-3
Sample IV.....	A-3
Sample V.....	A-4
Sample VI.....	A-4
Sample VII.....	A-5
Sample VIII.....	A-5
Sample IX.....	A-6

Column set 2:

Sample I.....	A-6
Sample II.....	A-7
Sample III.....	A-7
Sample IV.....	A-8
Sample V.....	A-8
Sample VI.....	A-9
Sample VII.....	A-9
Sample VIII.....	A-10
Sample IX.....	A-10

Column set 3:

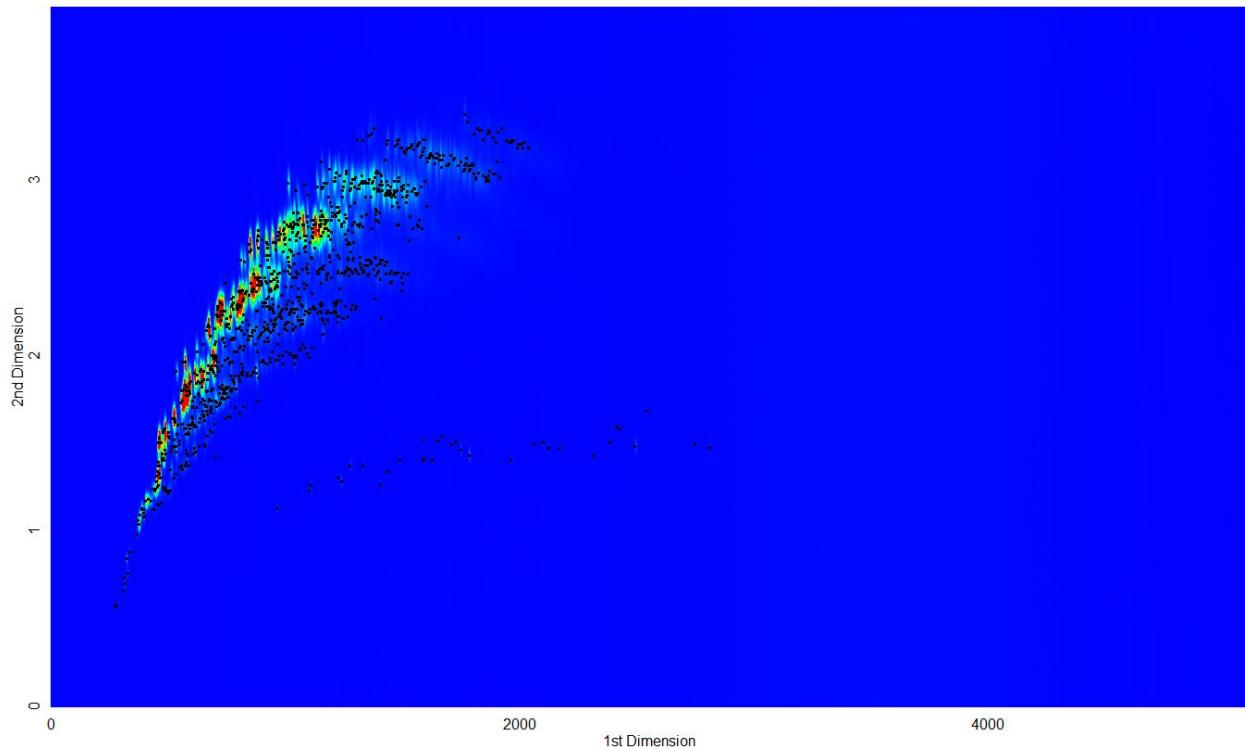
Sample I.....	A-11
Sample II.....	A-11
Sample III.....	A-12
Sample IV.....	A-12
Sample V.....	A-13
Sample VI.....	A-13
Sample VII.....	A-14
Sample VIII.....	A-14
Sample IX.....	A-15

Column set 4:

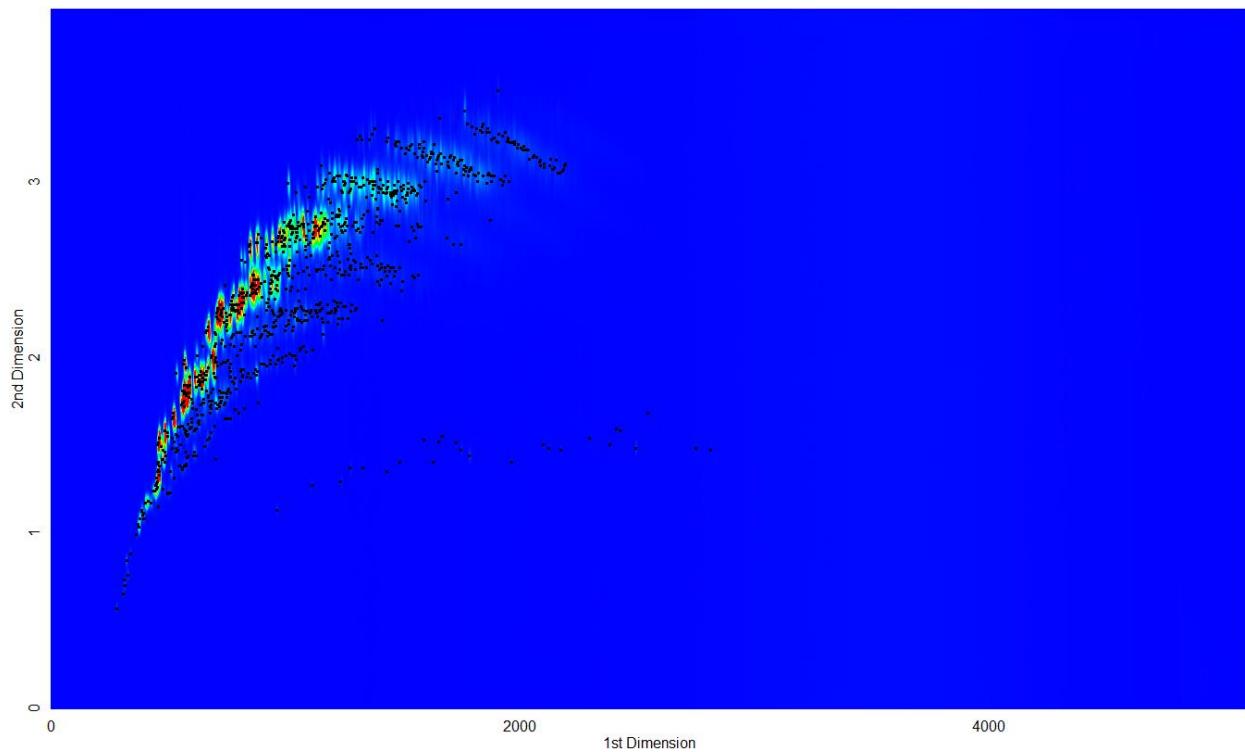
Sample I.....	A-15
Sample II.....	A-16
Sample III.....	A-16
Sample IV.....	A-17
Sample V.....	A-17
Sample VI.....	A-18
Sample VII.....	A-18
Sample VIII.....	A-19
Sample IX.....	A-19

Note that retention times are reported in seconds for both dimensions of all samples.

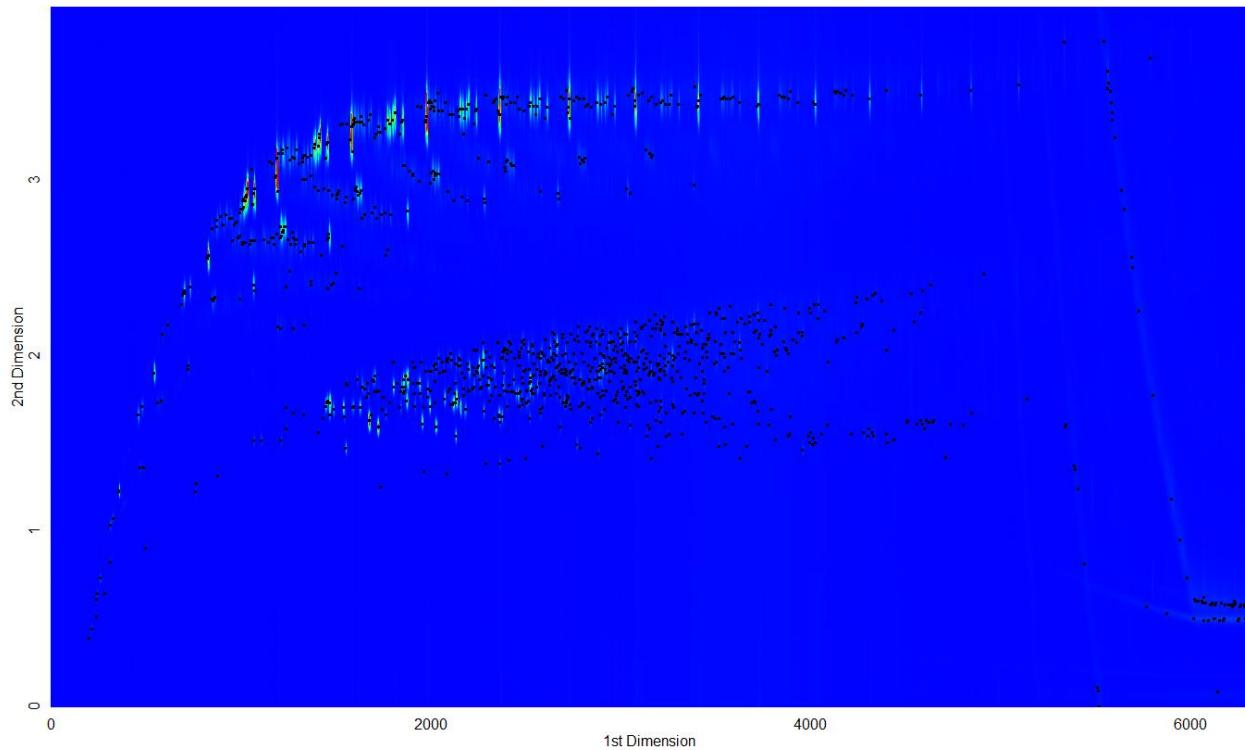
Column set 1: Sample I



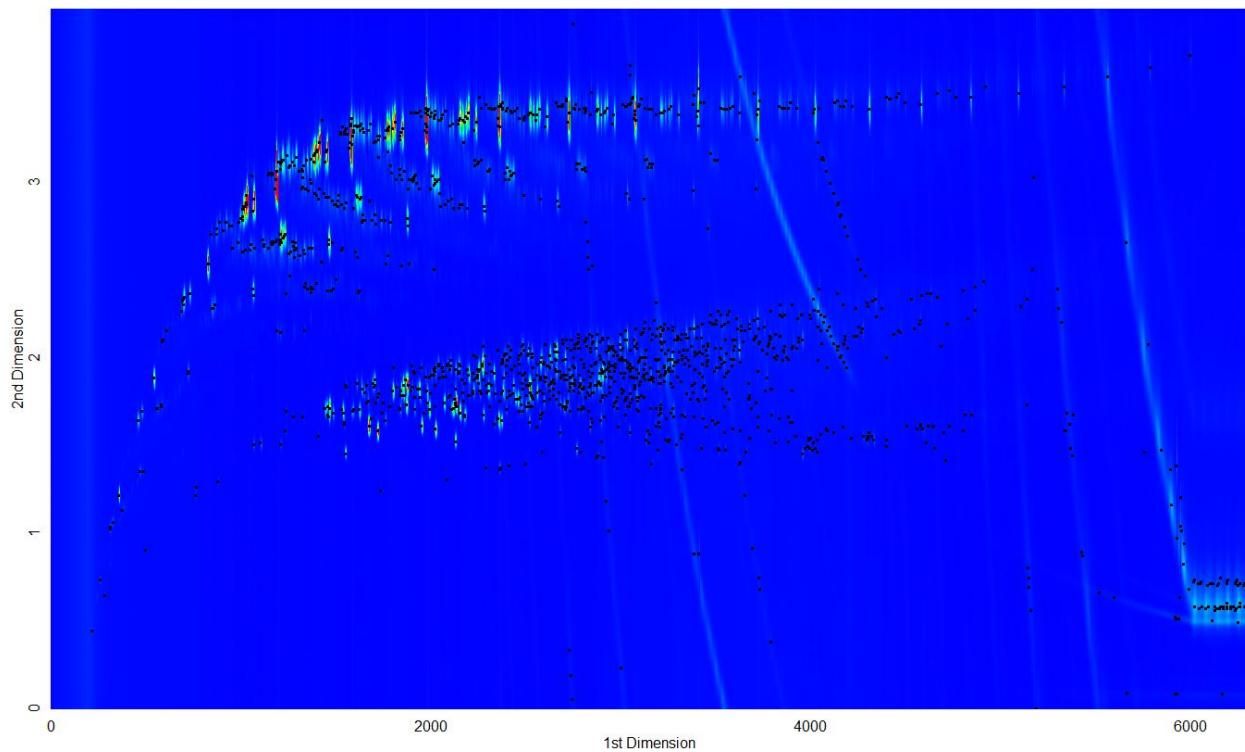
Column set 1: Sample II



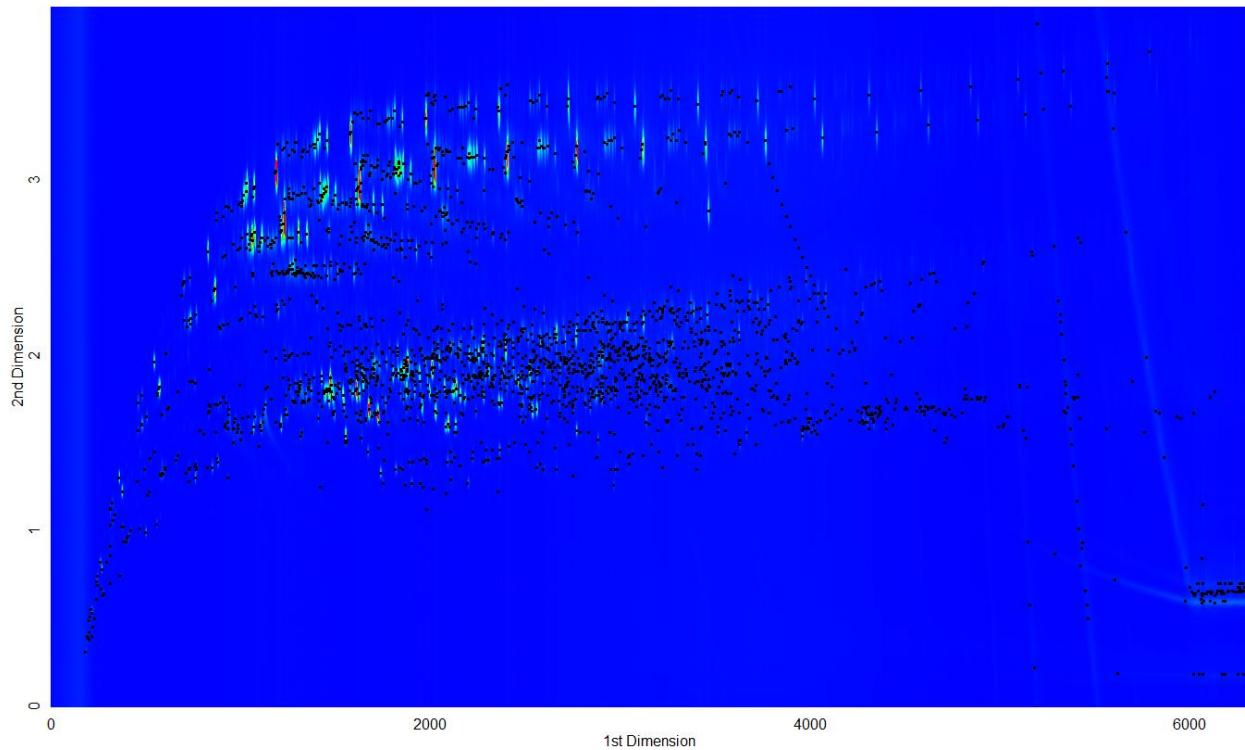
Column set 1: Sample III



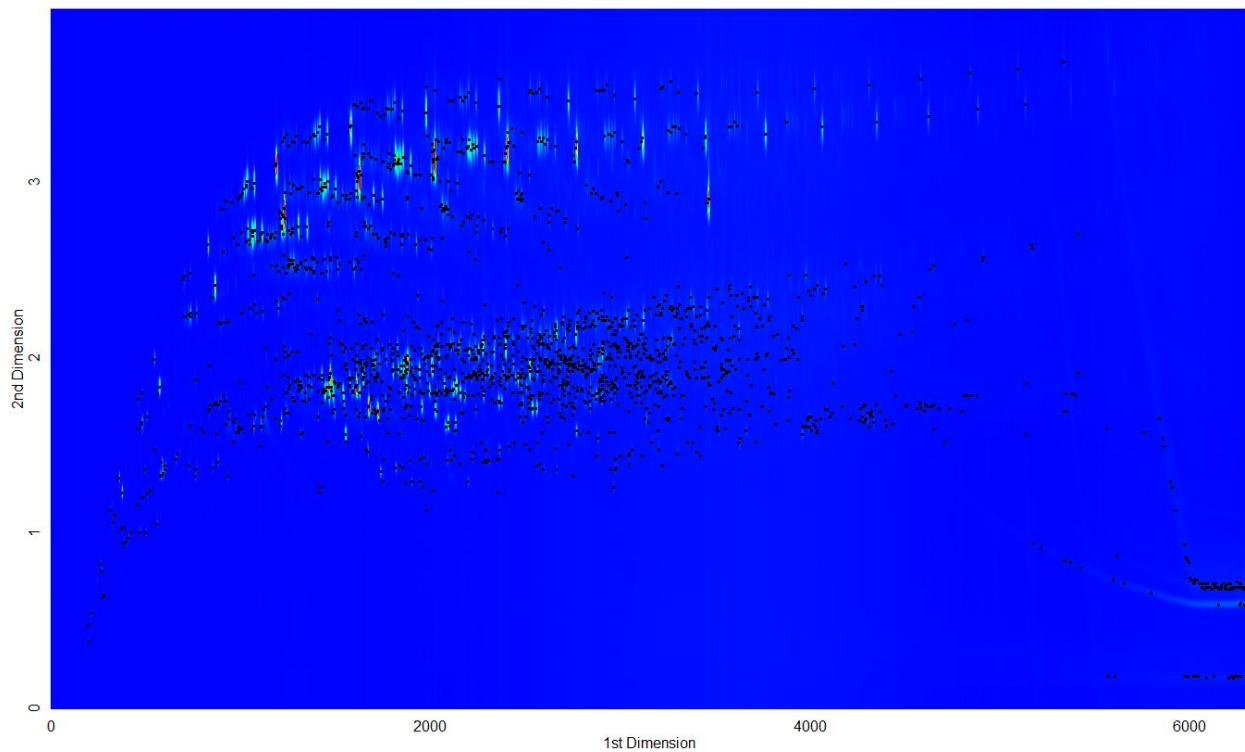
Column set 1: Sample IV



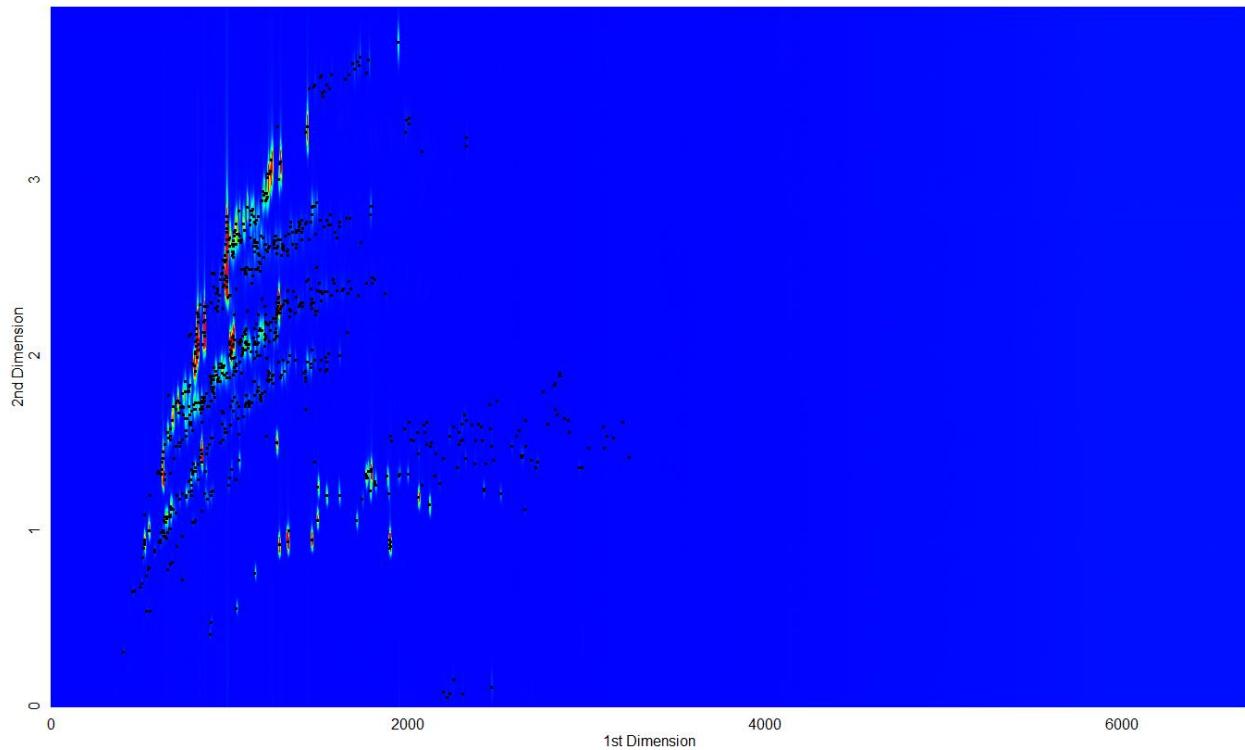
Column set 1: Sample V



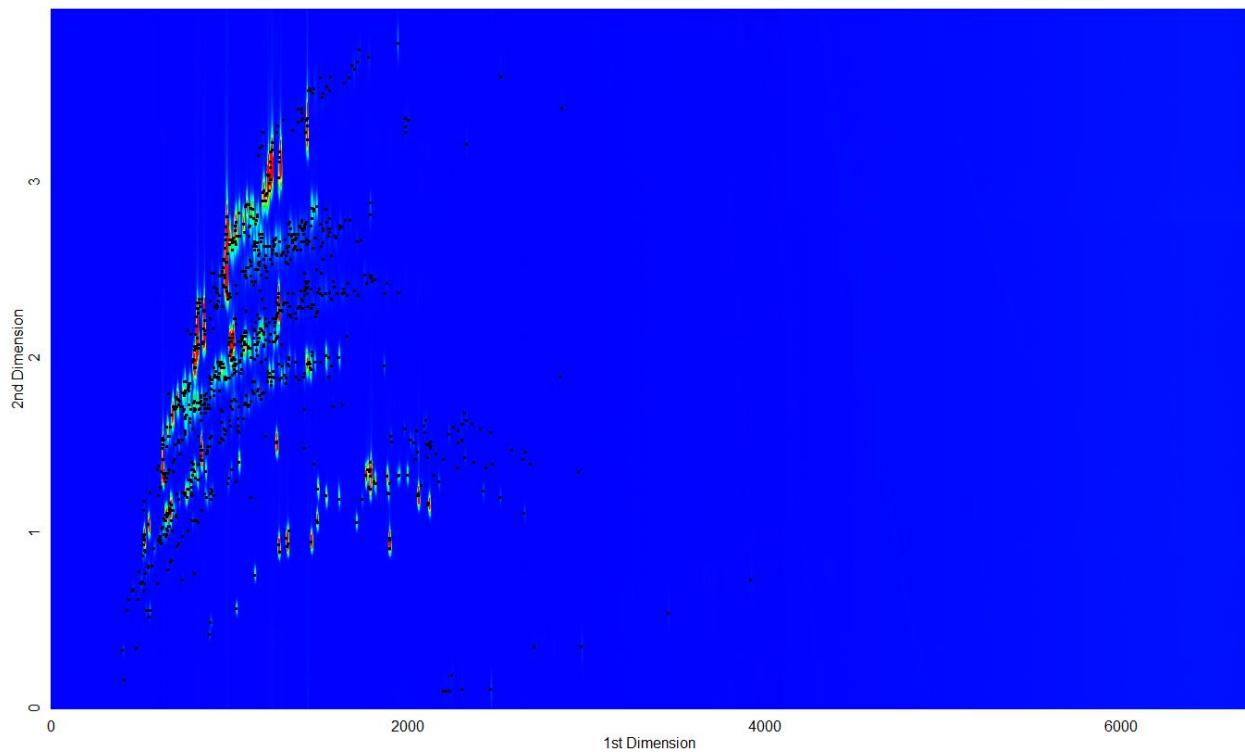
Column set 1: Sample VI



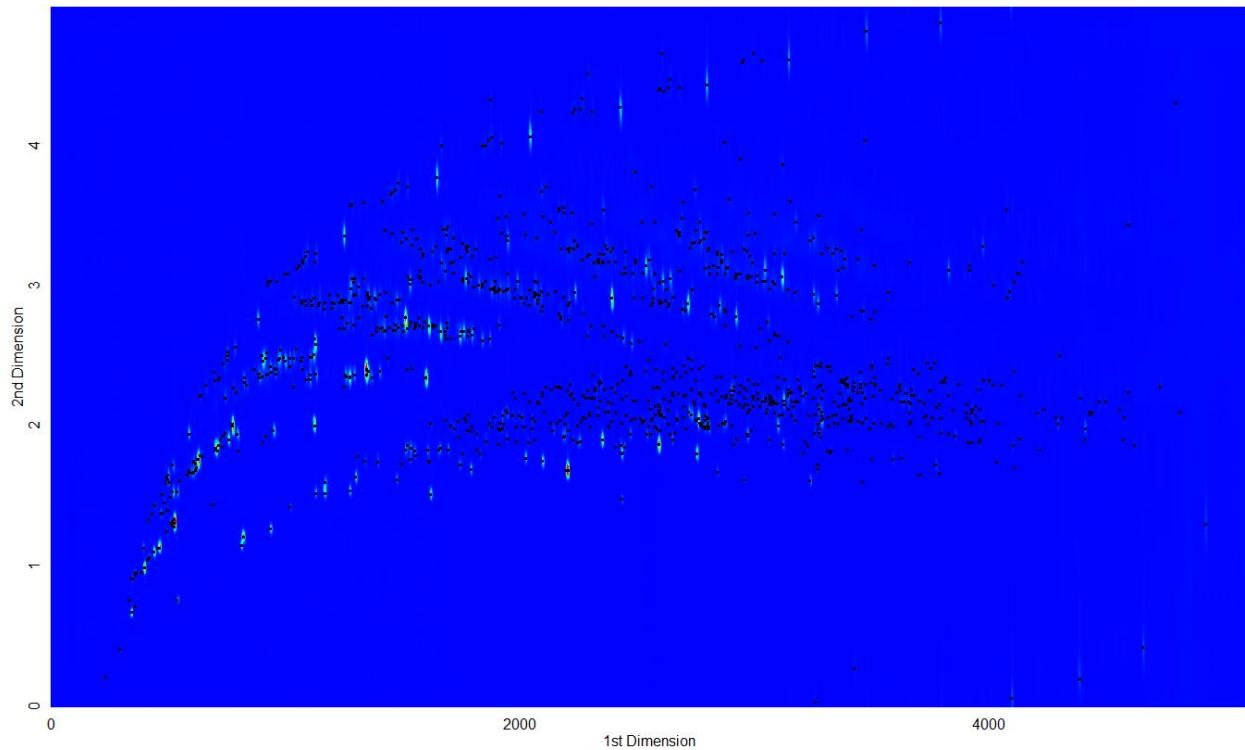
Column set 1: Sample VII



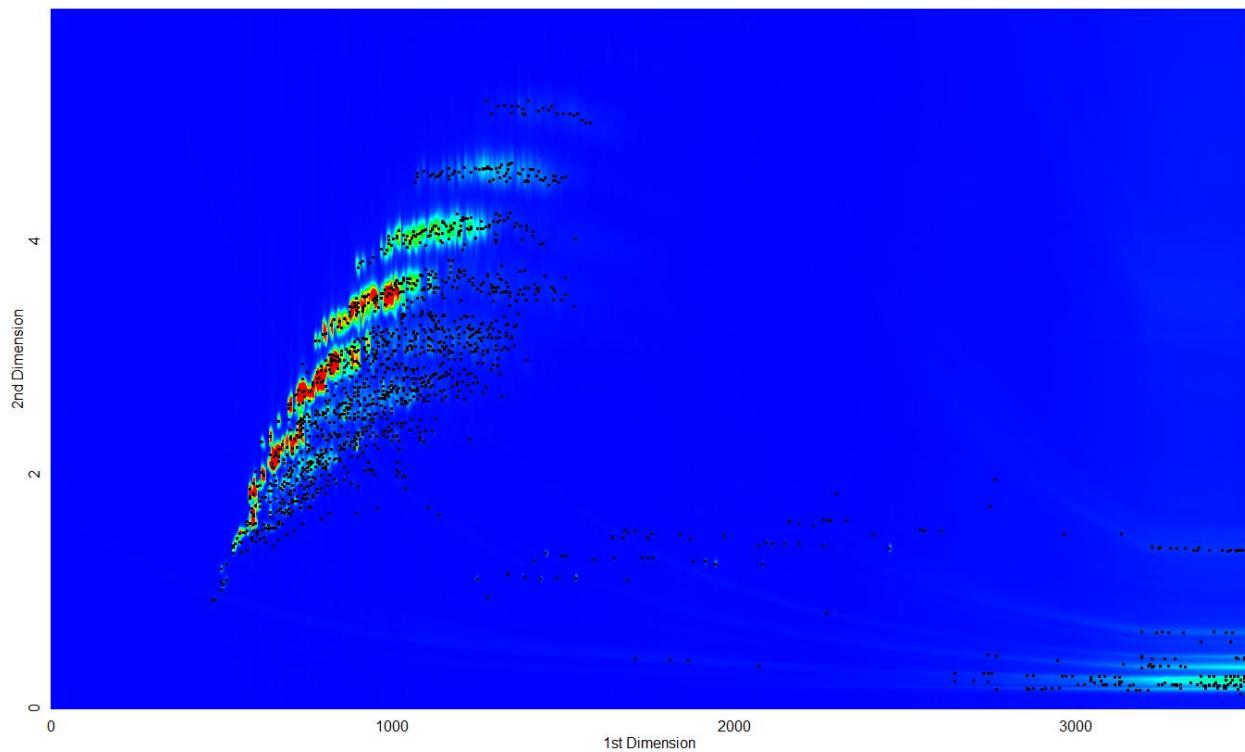
Column set 1: Sample VIII



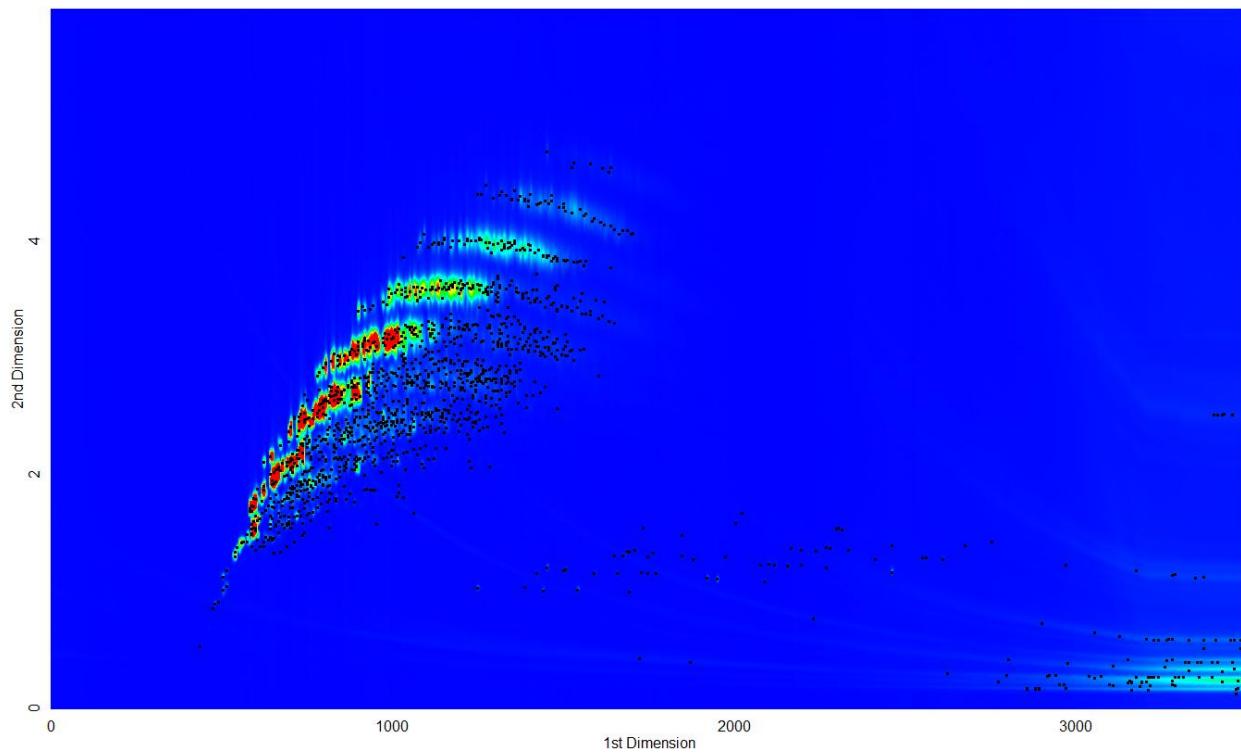
Column set 1: Sample IX



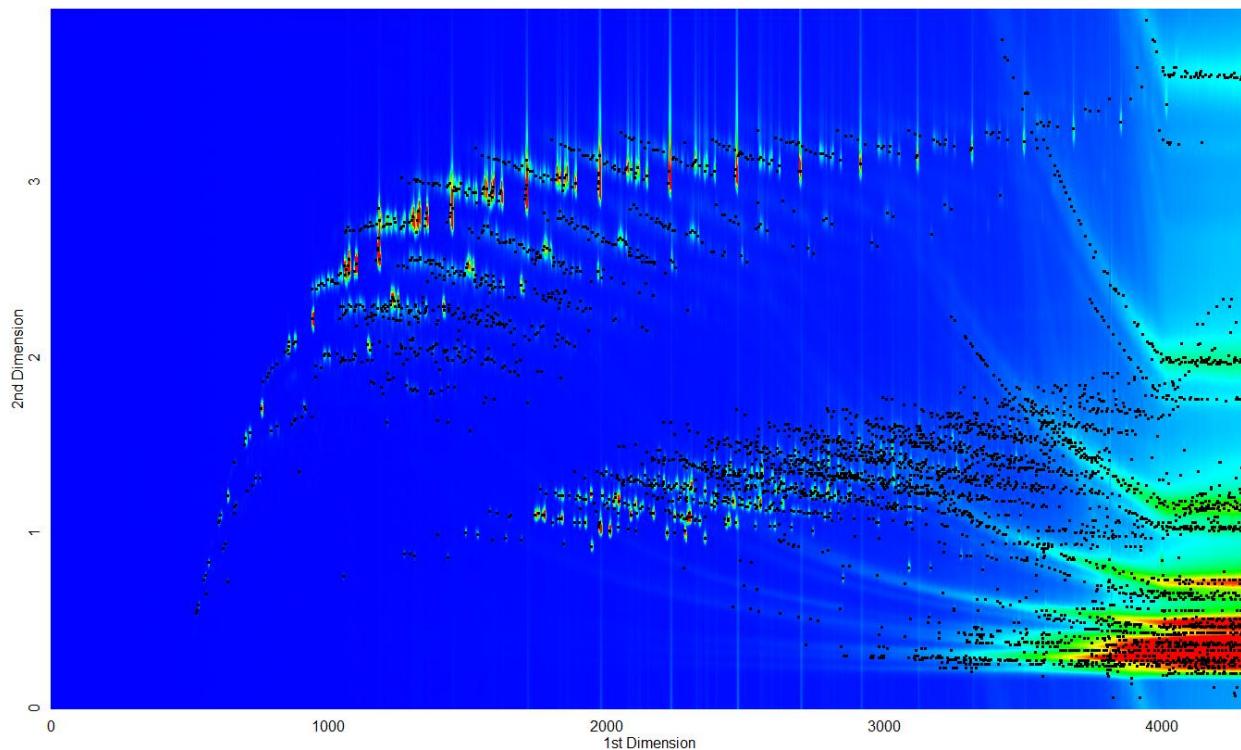
Column set 2: Sample I



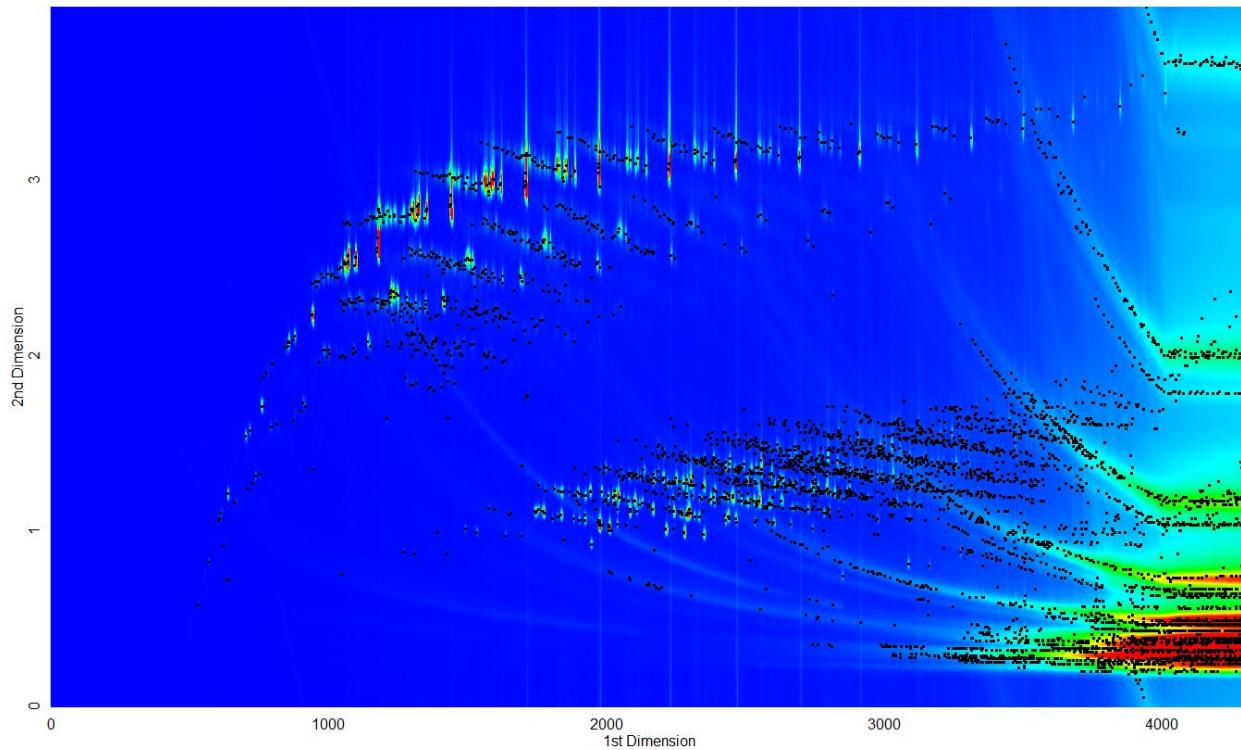
Column set 2: Sample II



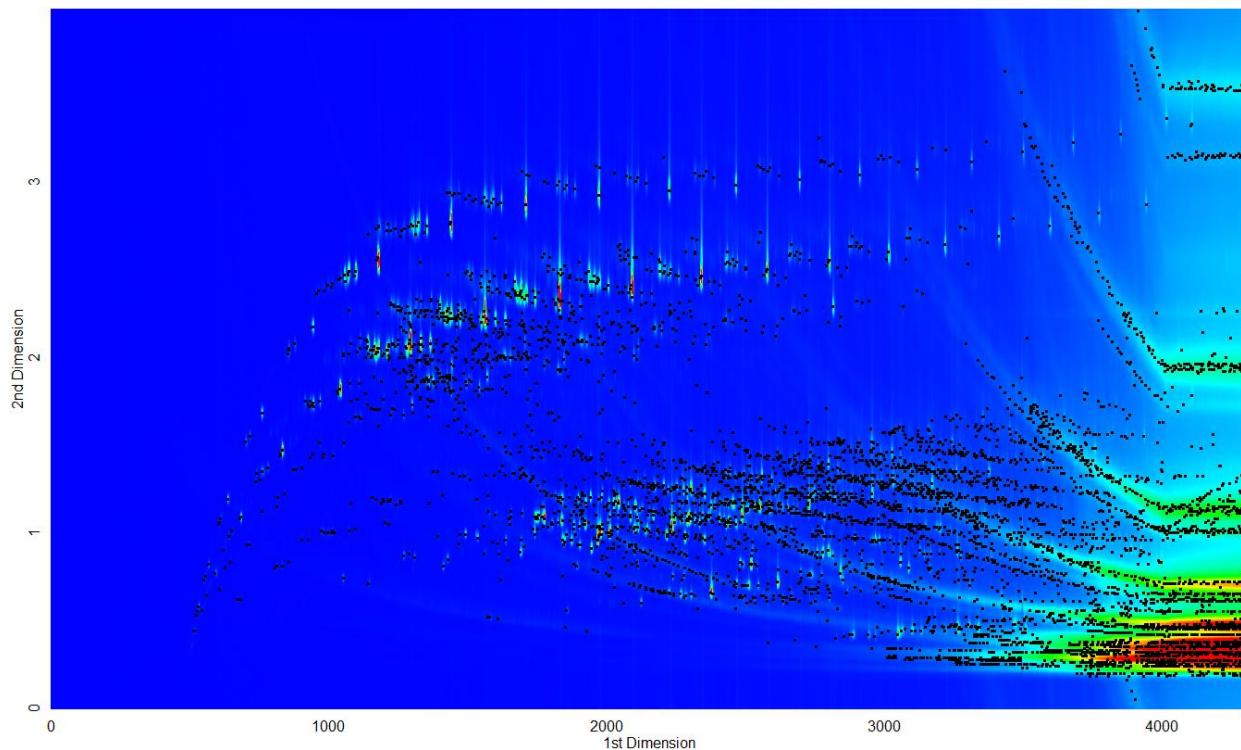
Column set 2: Sample III



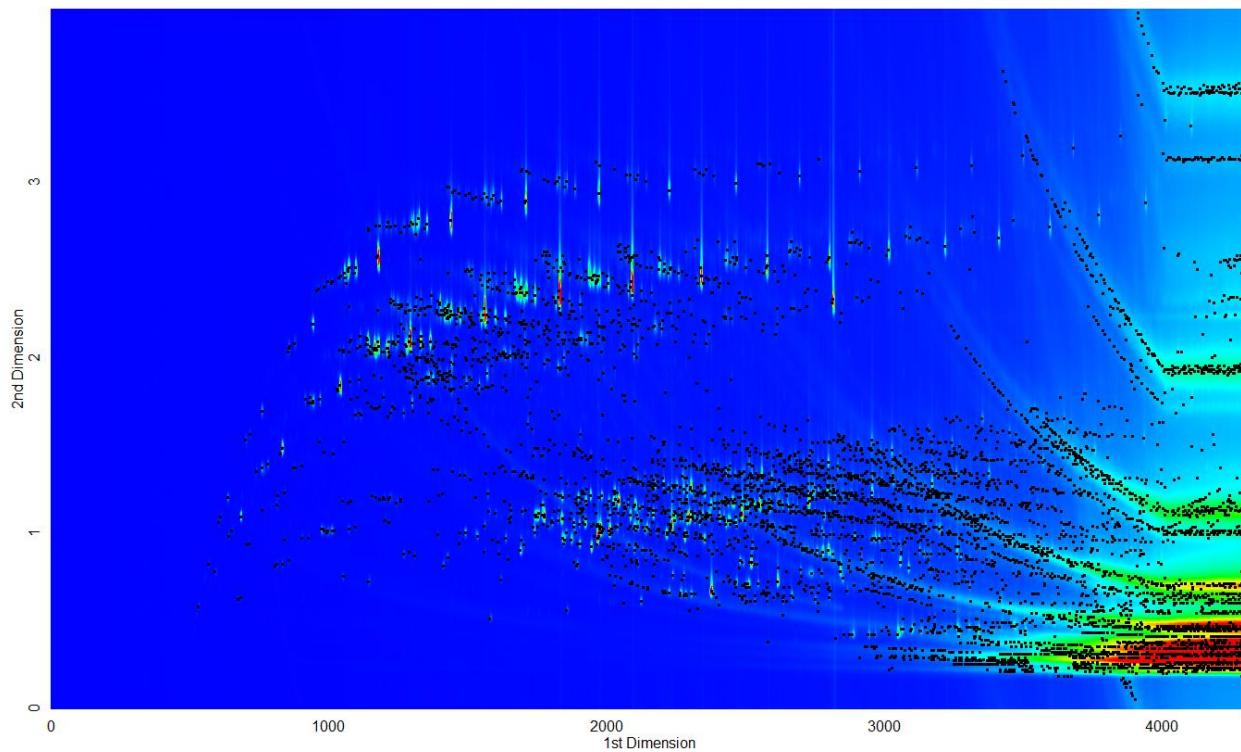
Column set 2: Sample IV



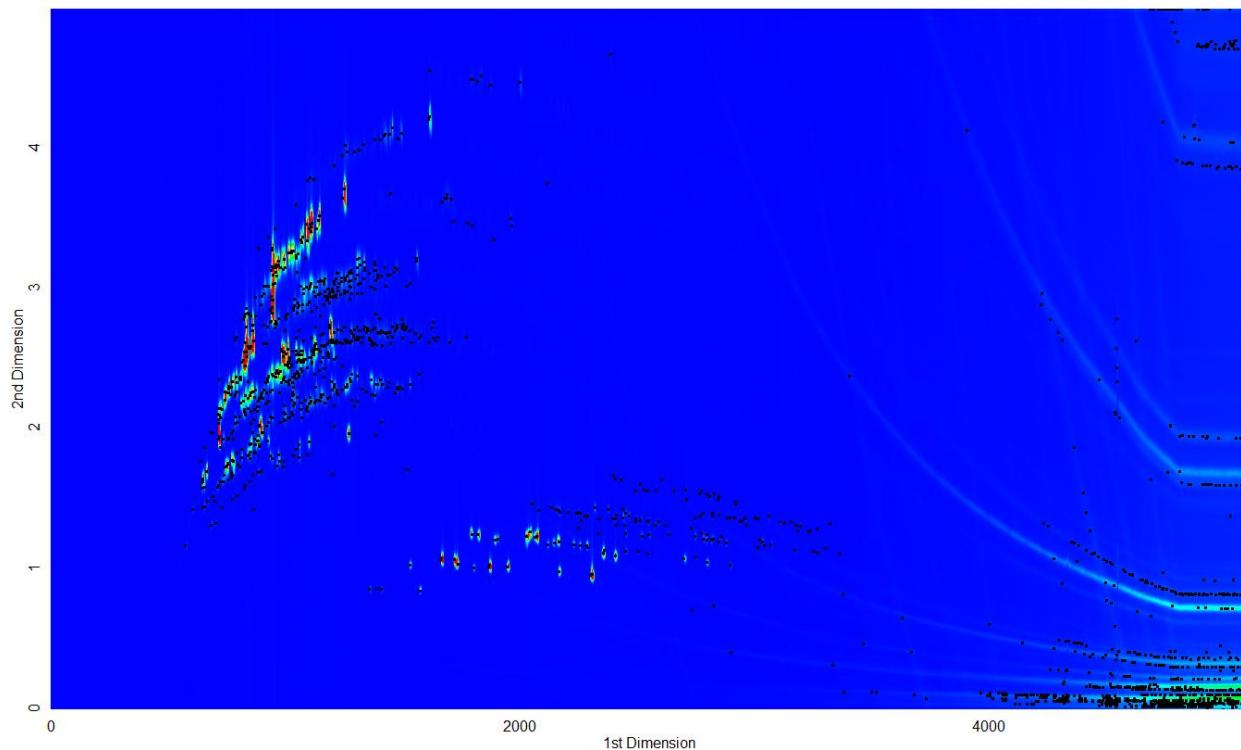
Column set 2: Sample V



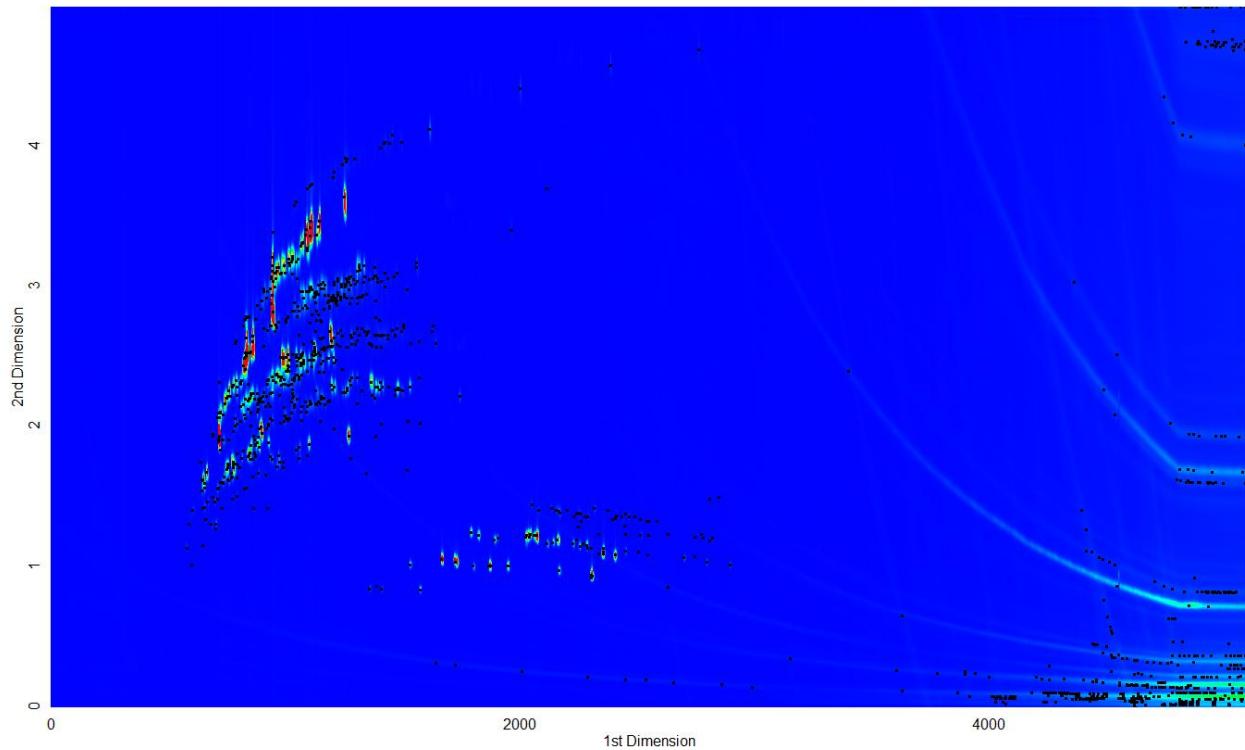
Column set 2: Sample VI



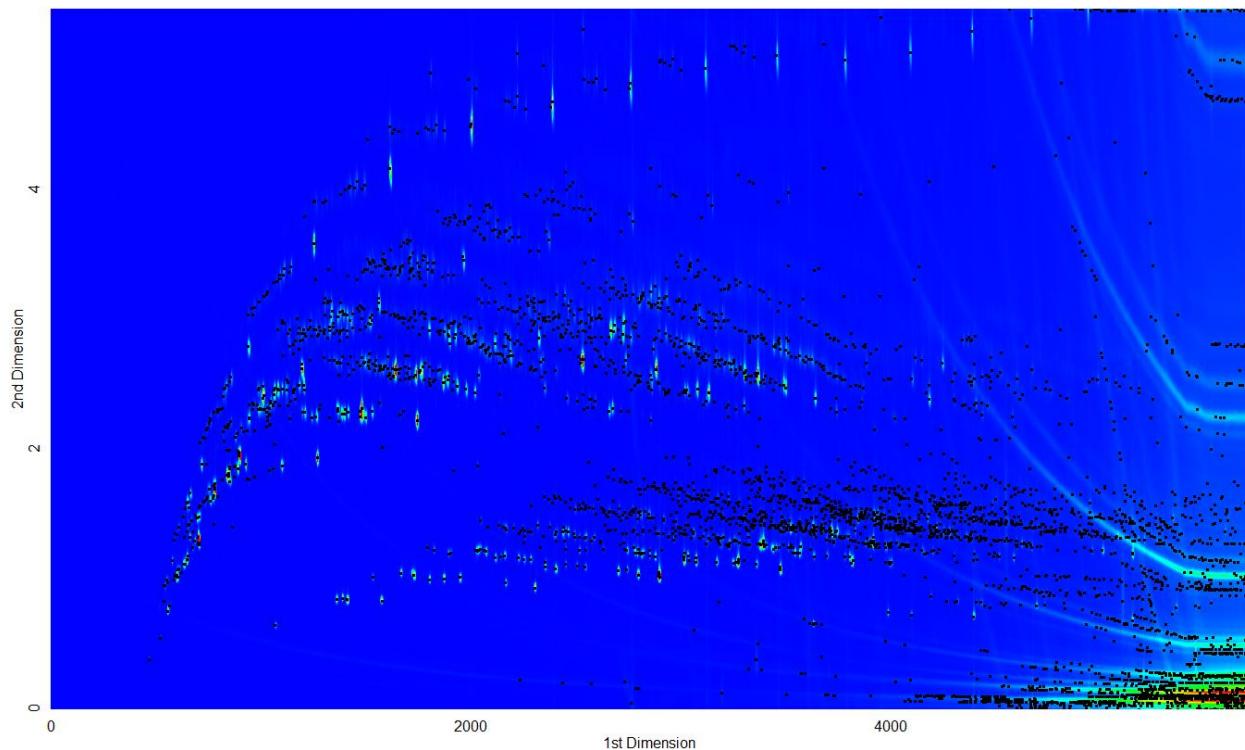
Column set 2: Sample VII



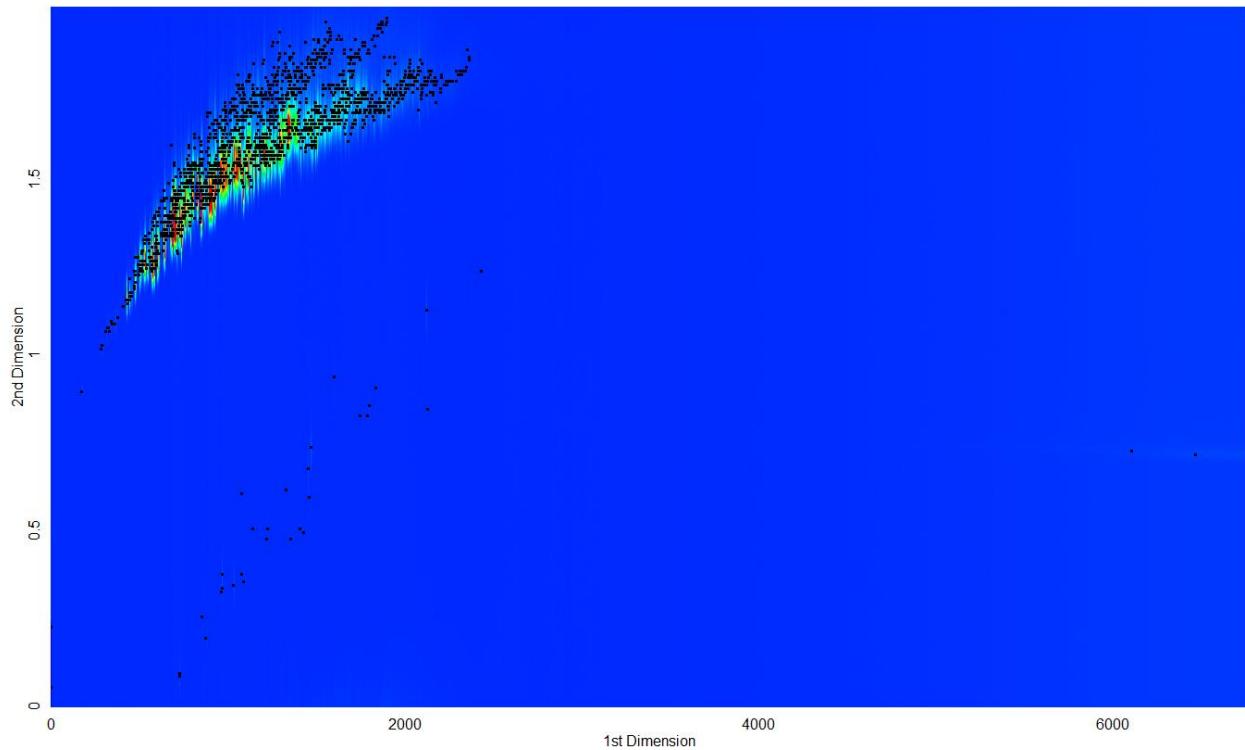
Column set 2: Sample VIII



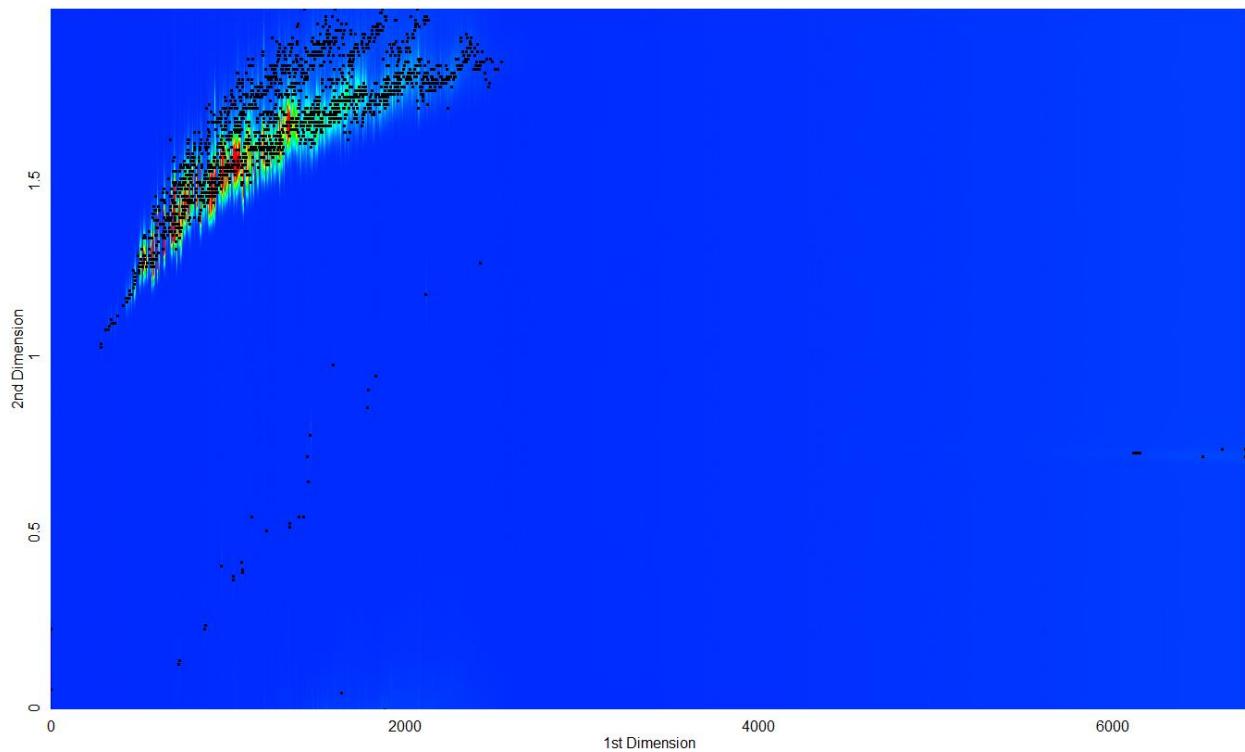
Column set 2: Sample IX



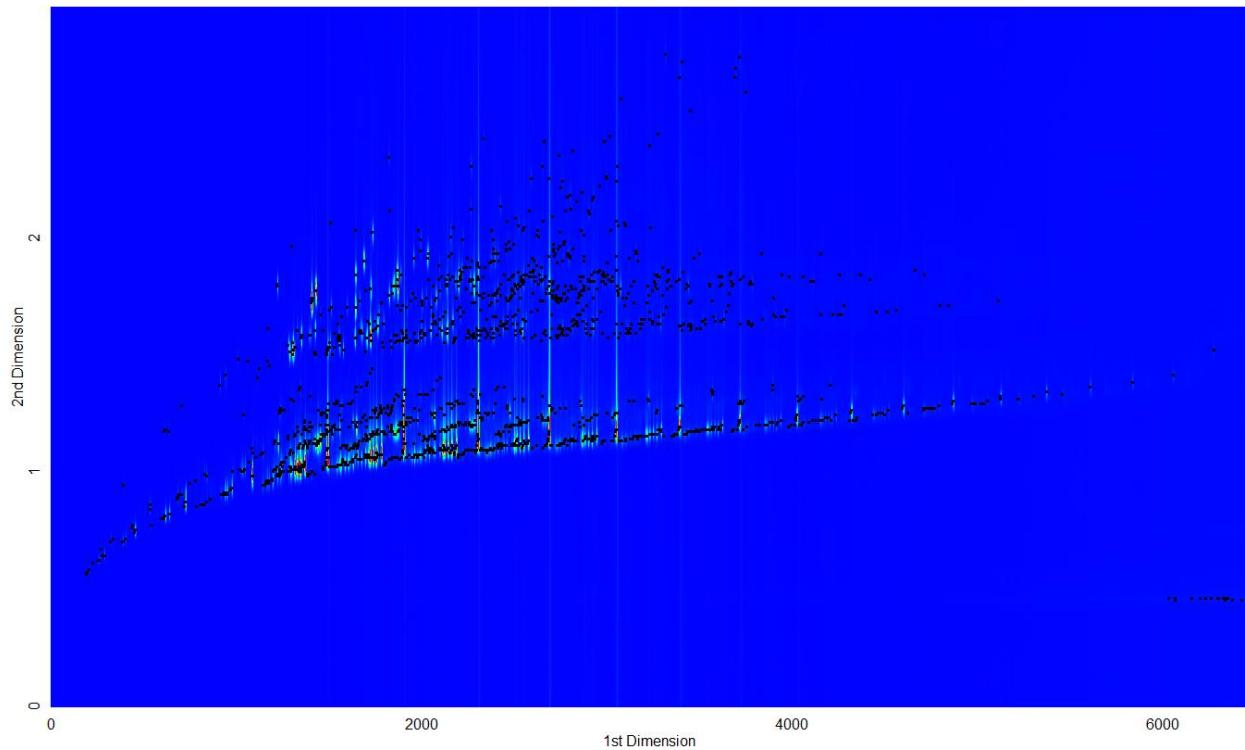
Column set 3: Sample I



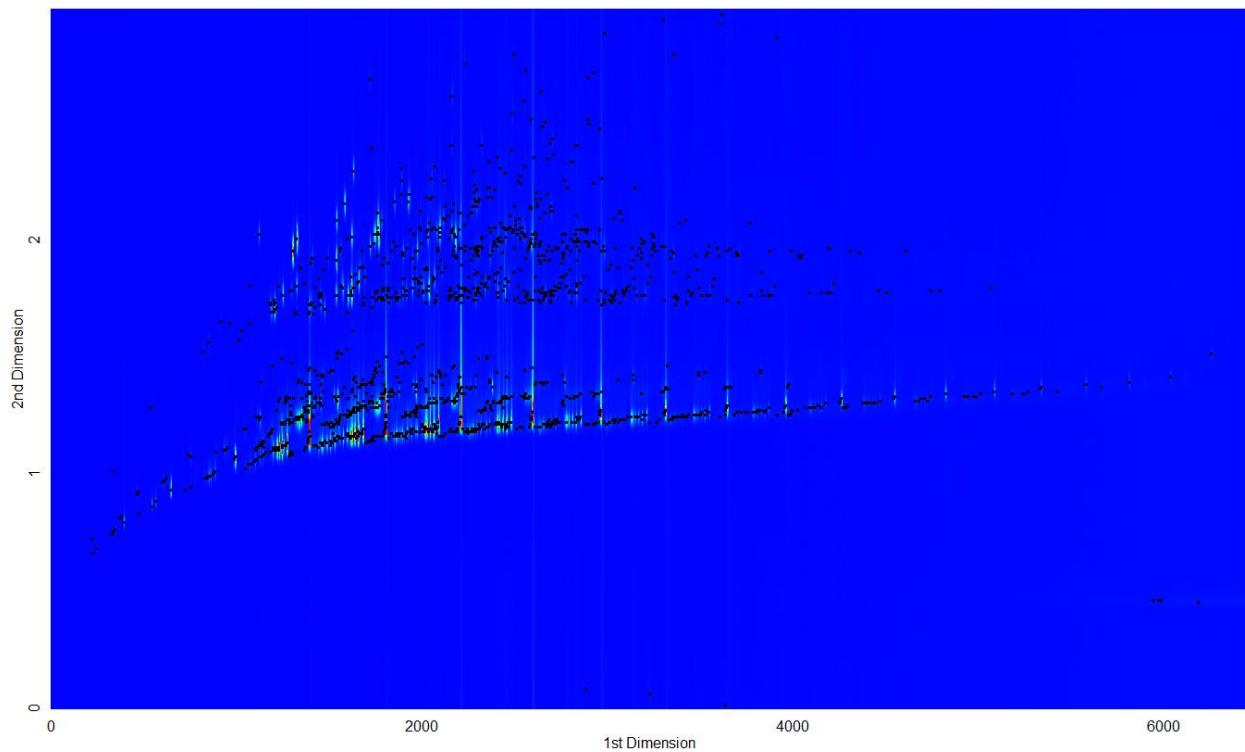
Column set 3: Sample II



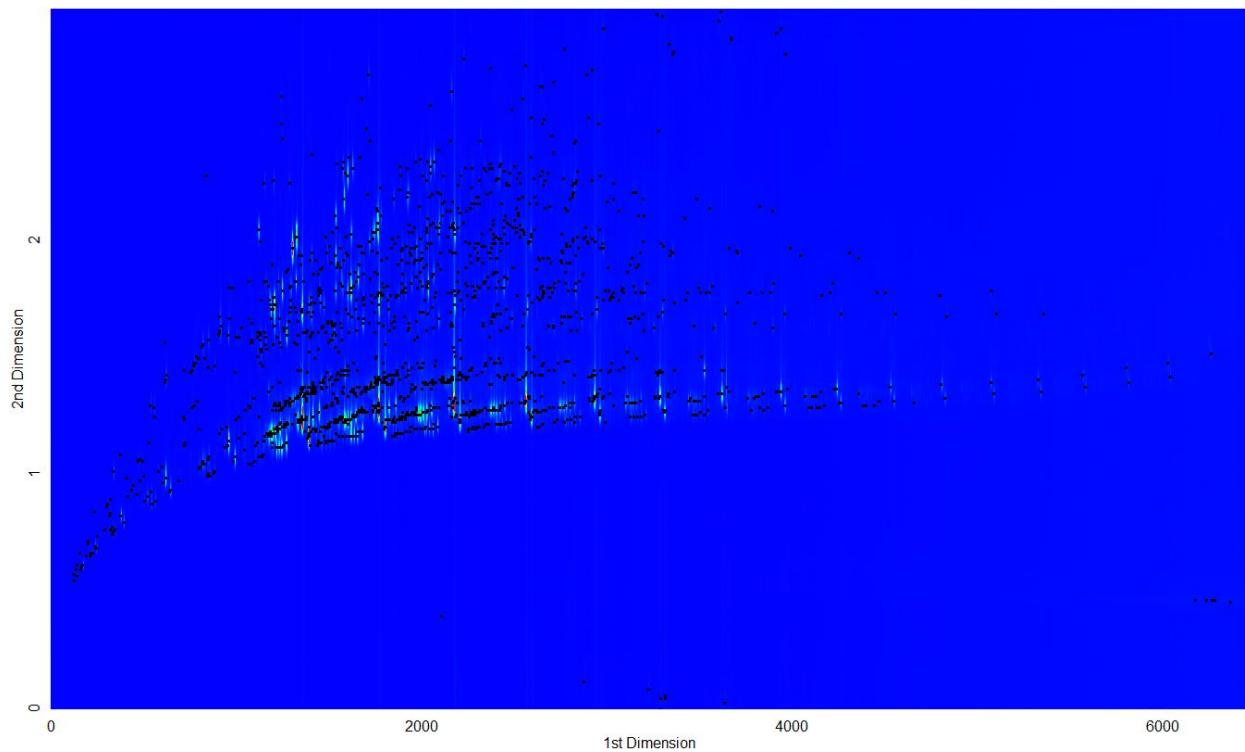
Column set 3: Sample III



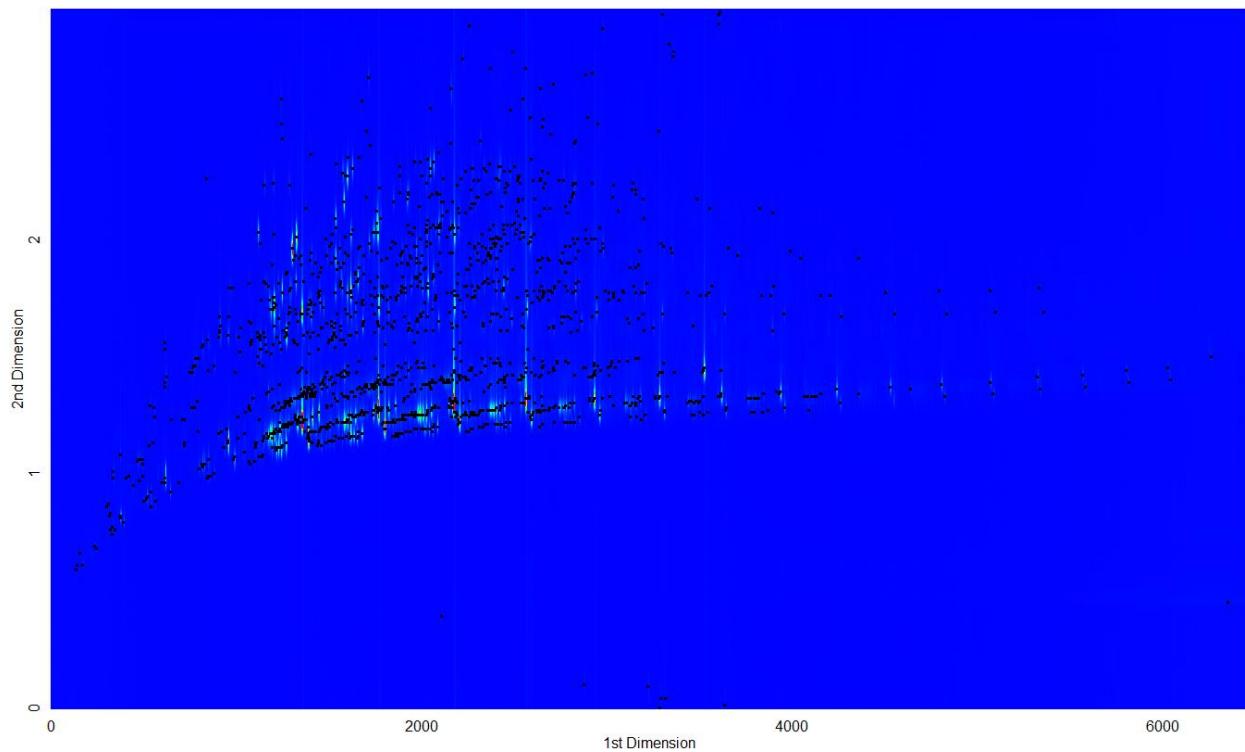
Column set 3: Sample IV



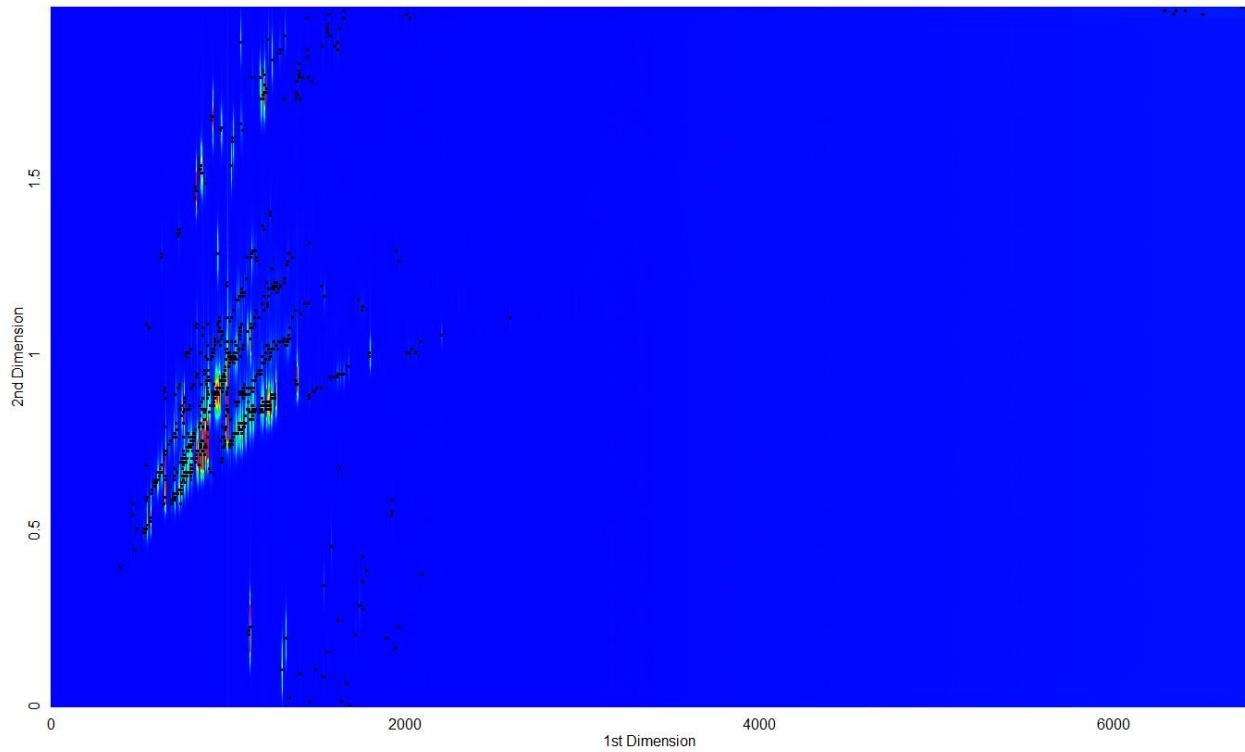
Column set 3: Sample V



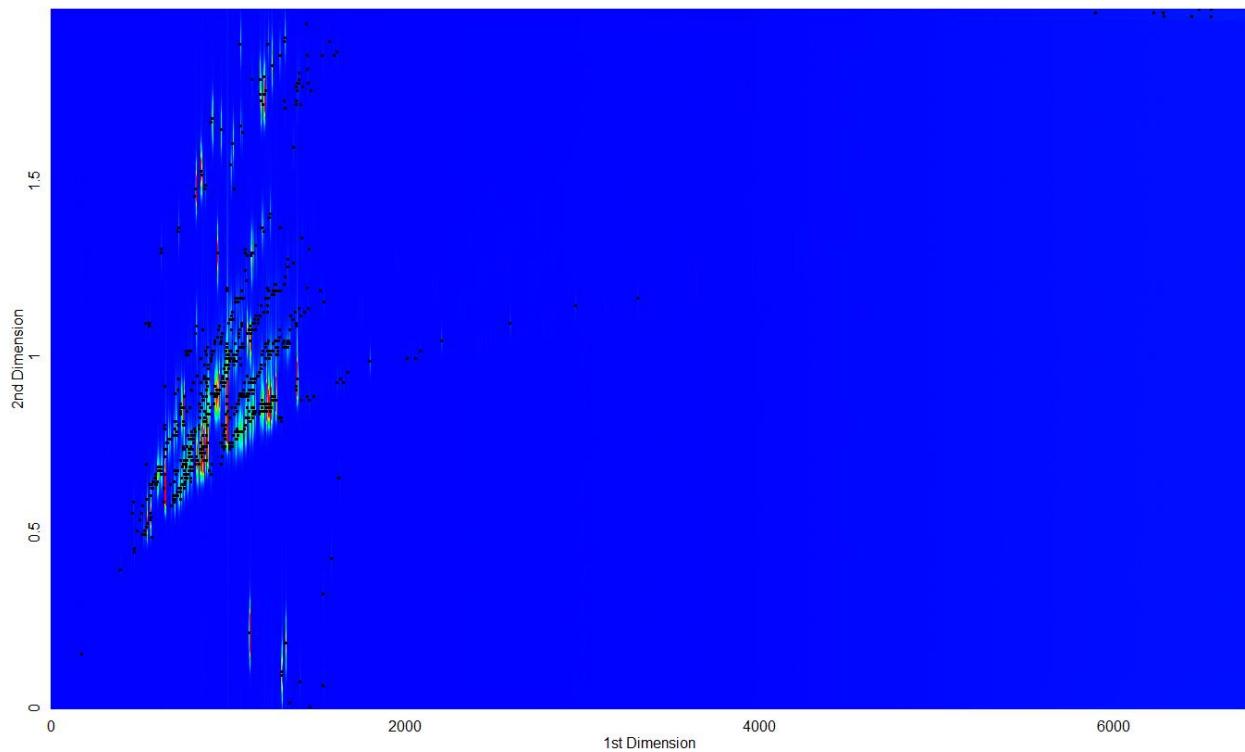
Column set 3: Sample VI



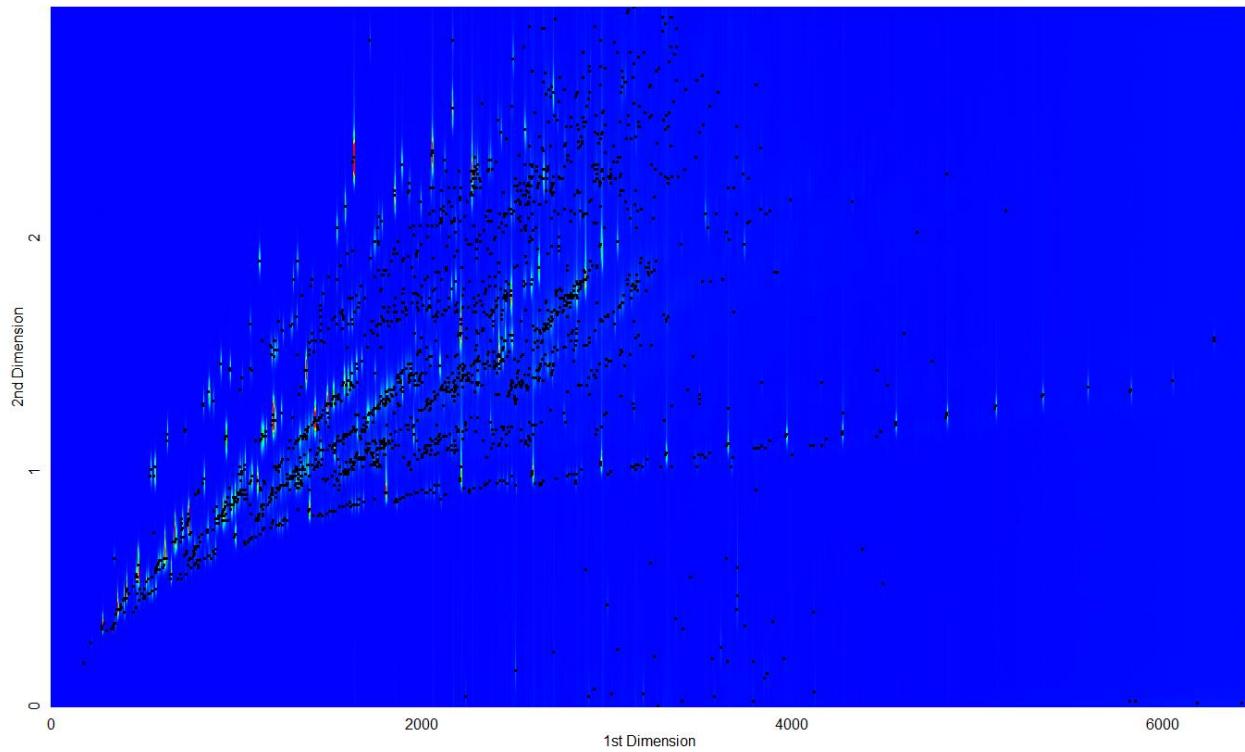
Column set 3: Sample VII



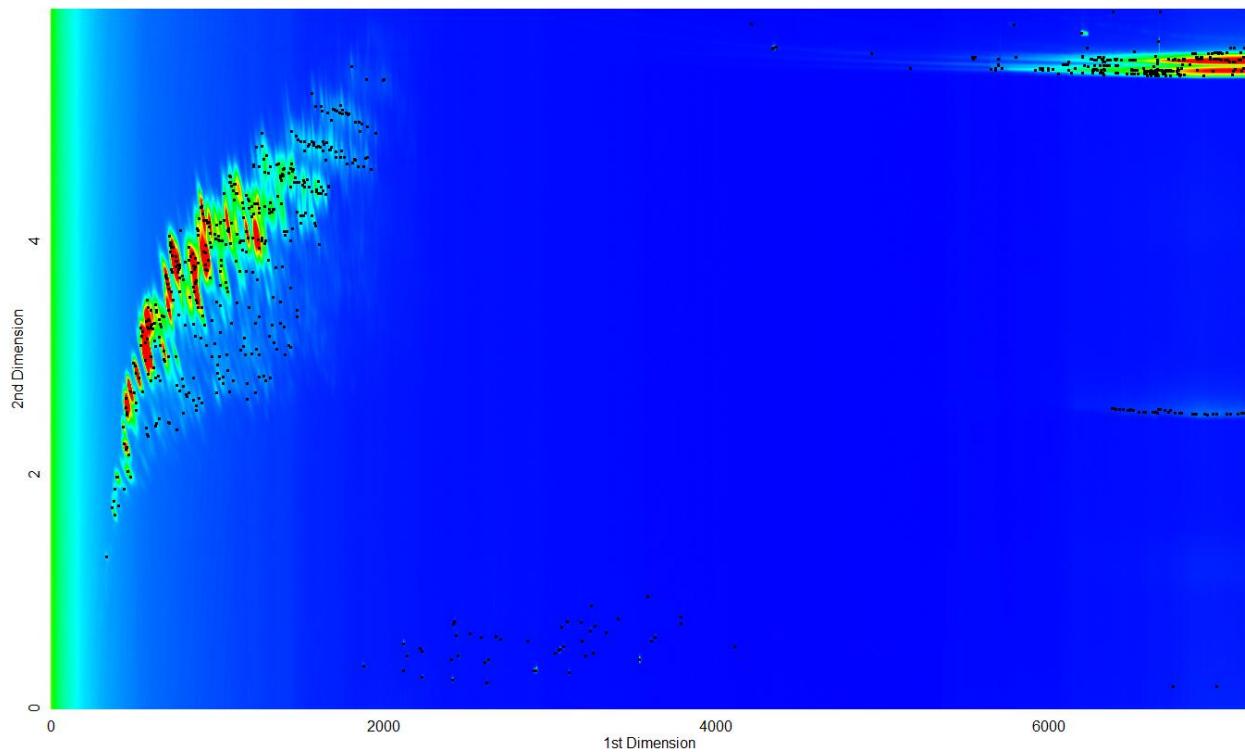
Column set 3: Sample VIII



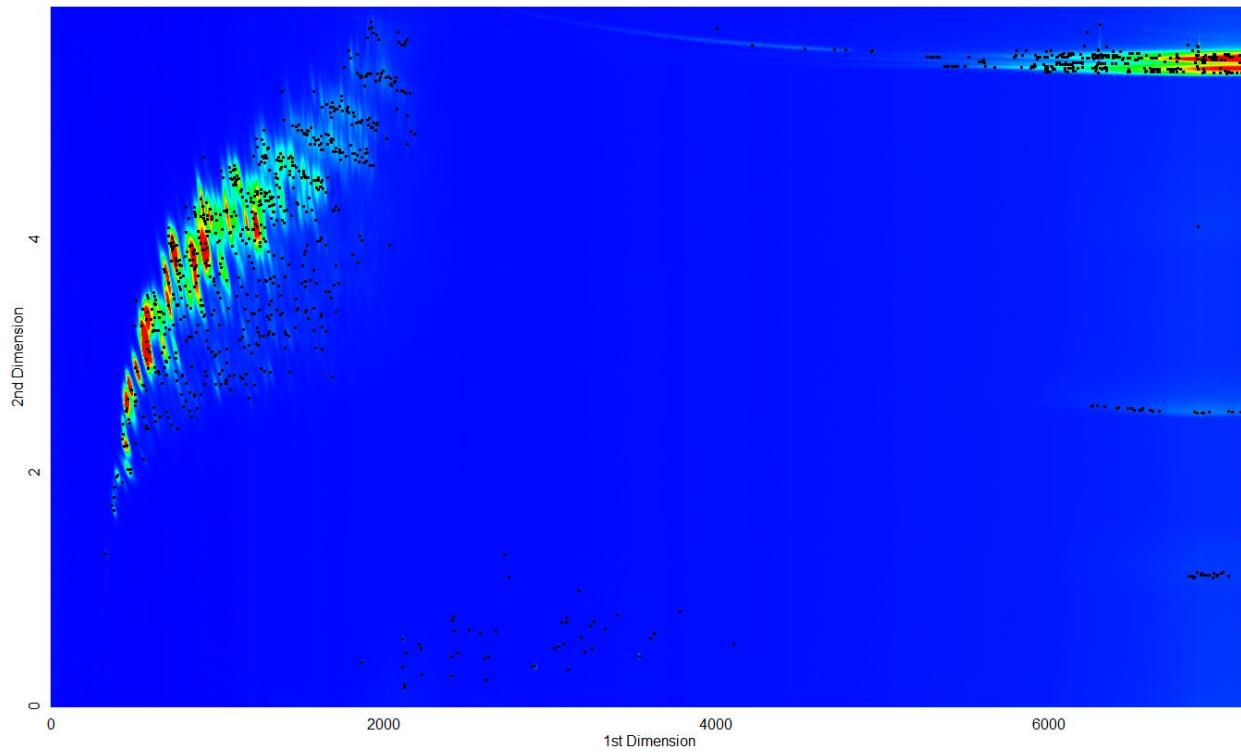
Column set 3: Sample IX



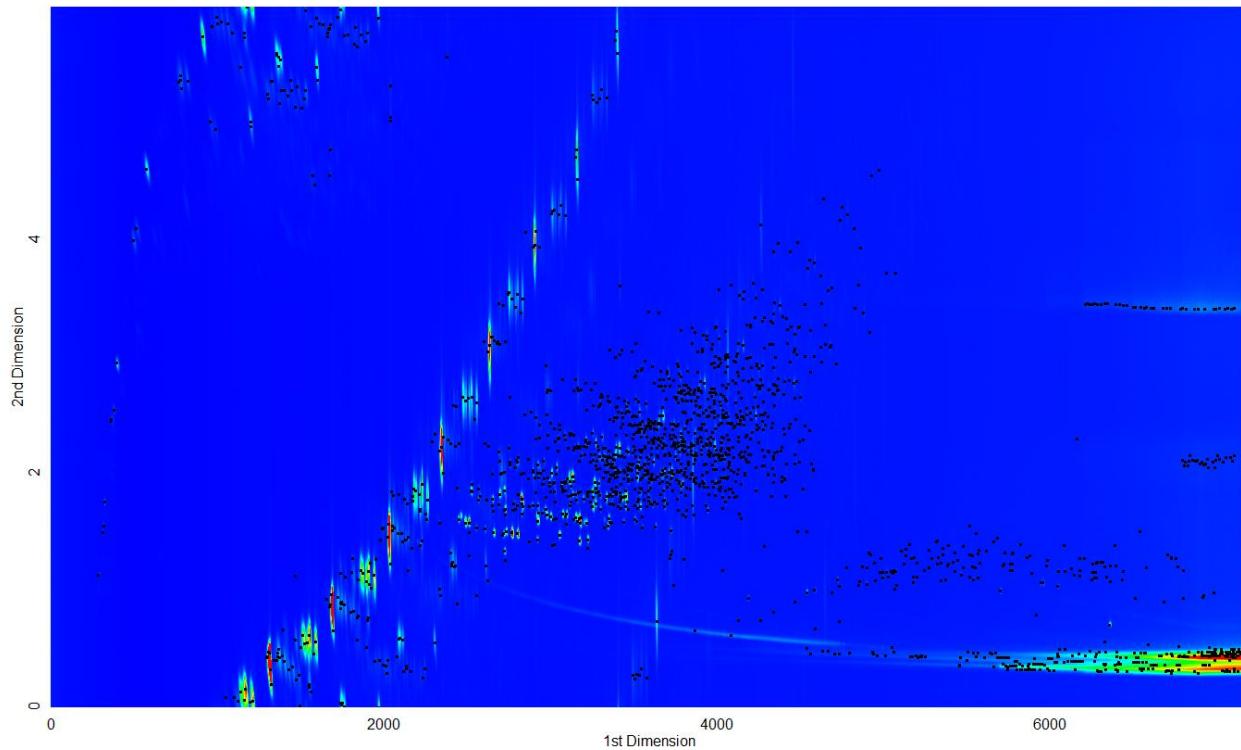
Column set 4: Sample I



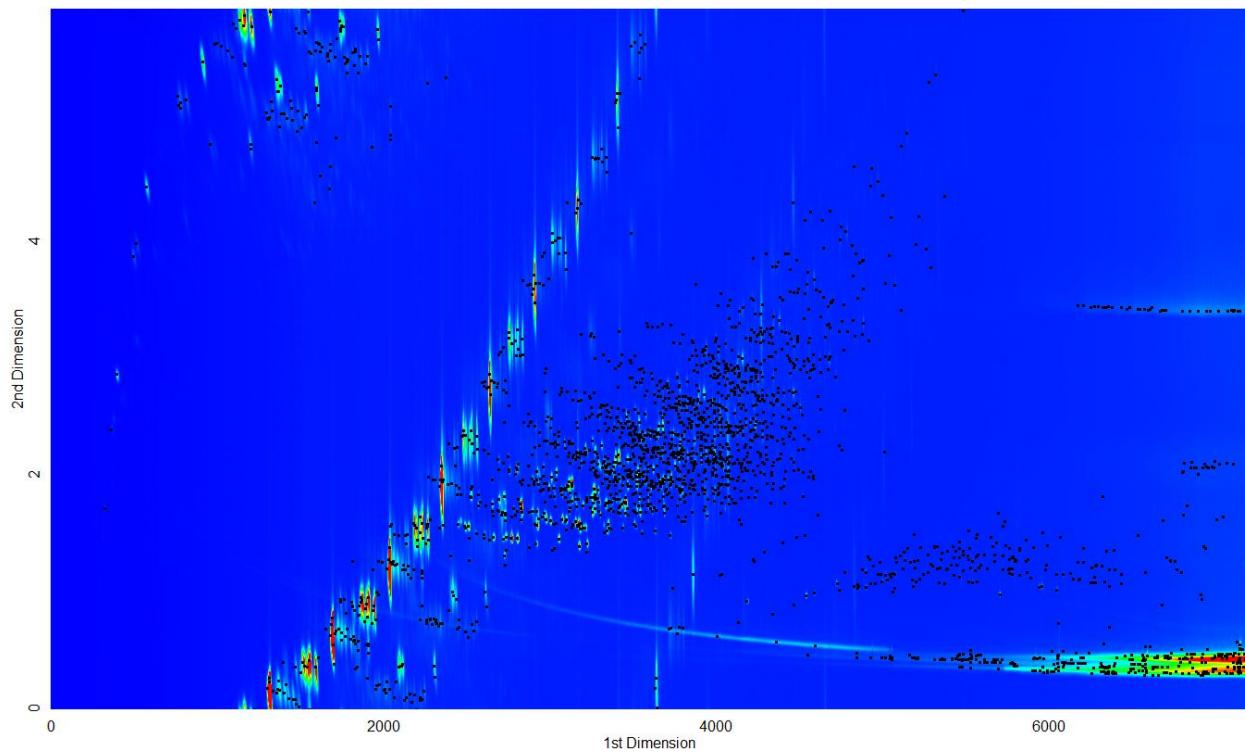
Column set 4: Sample II



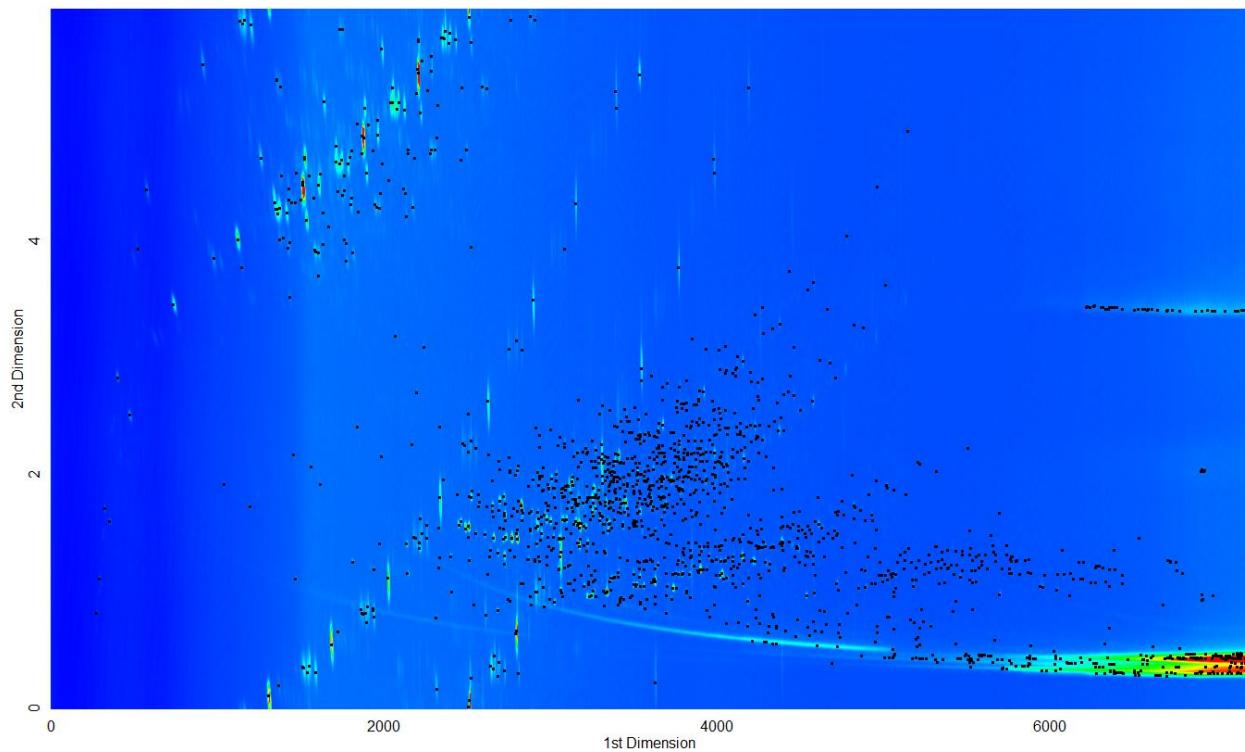
Column set 4: Sample III



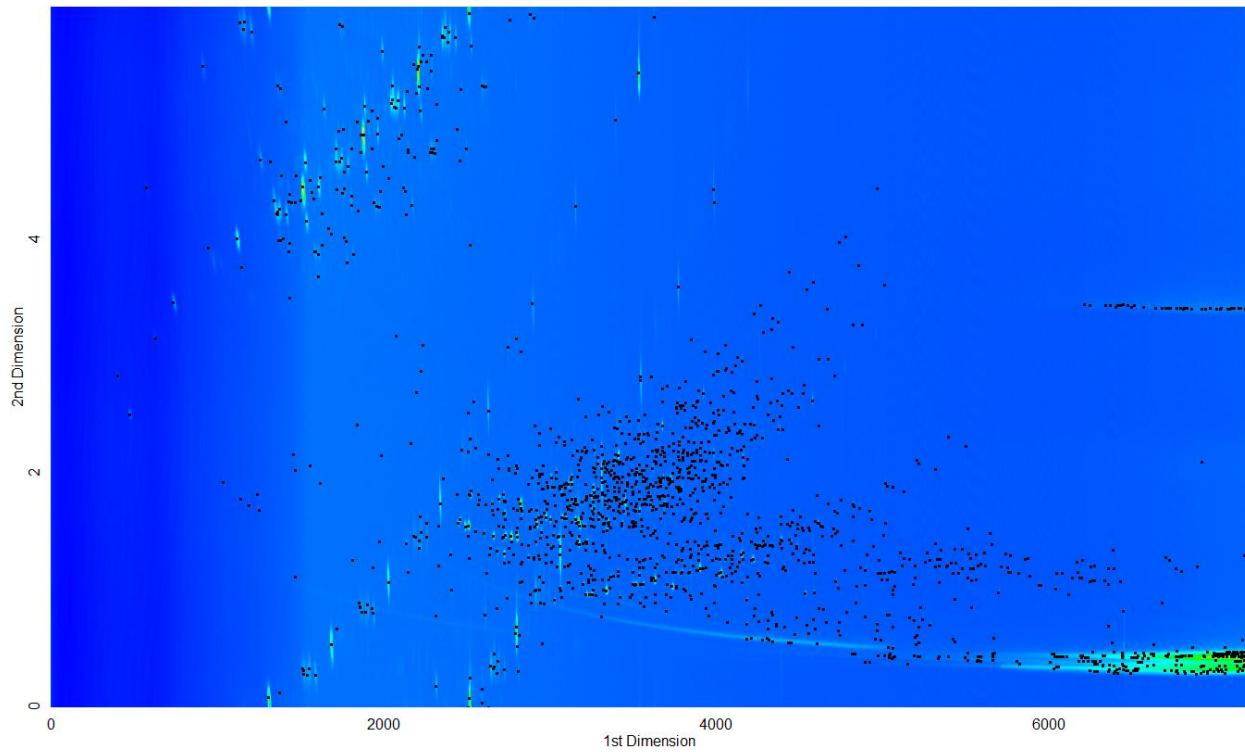
Column set 4: Sample IV



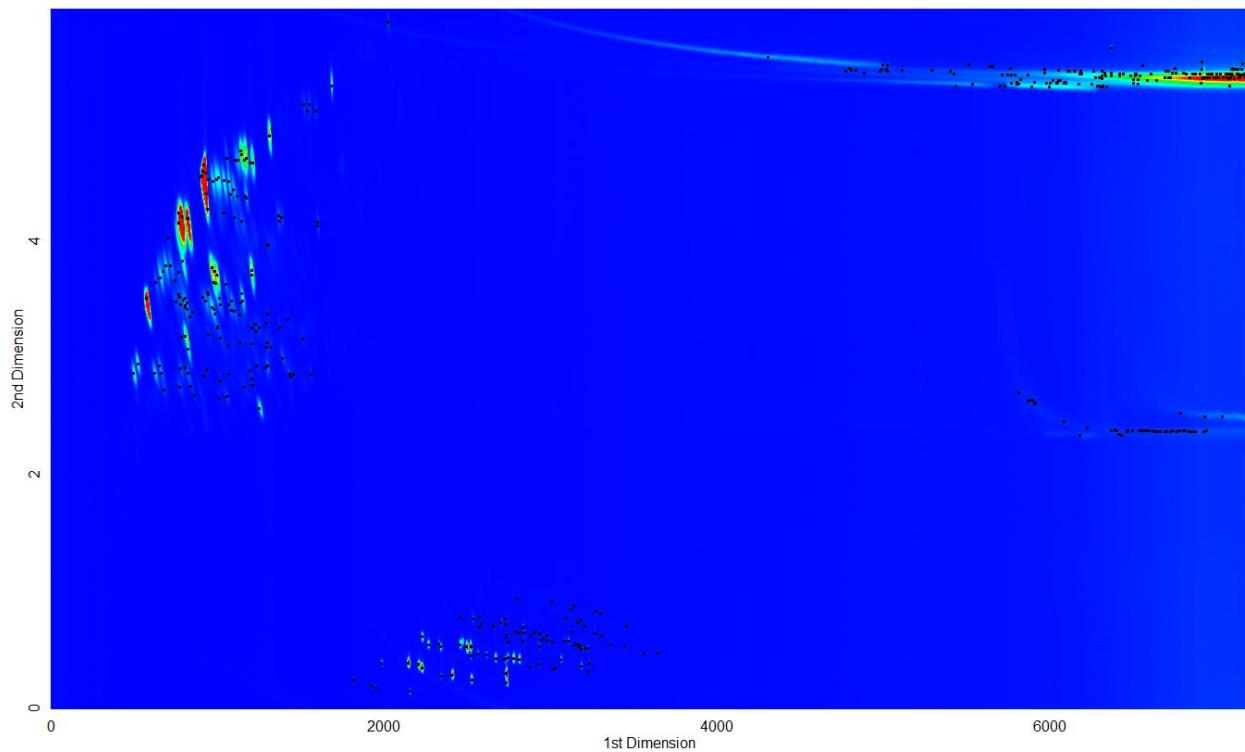
Column set 4: Sample V



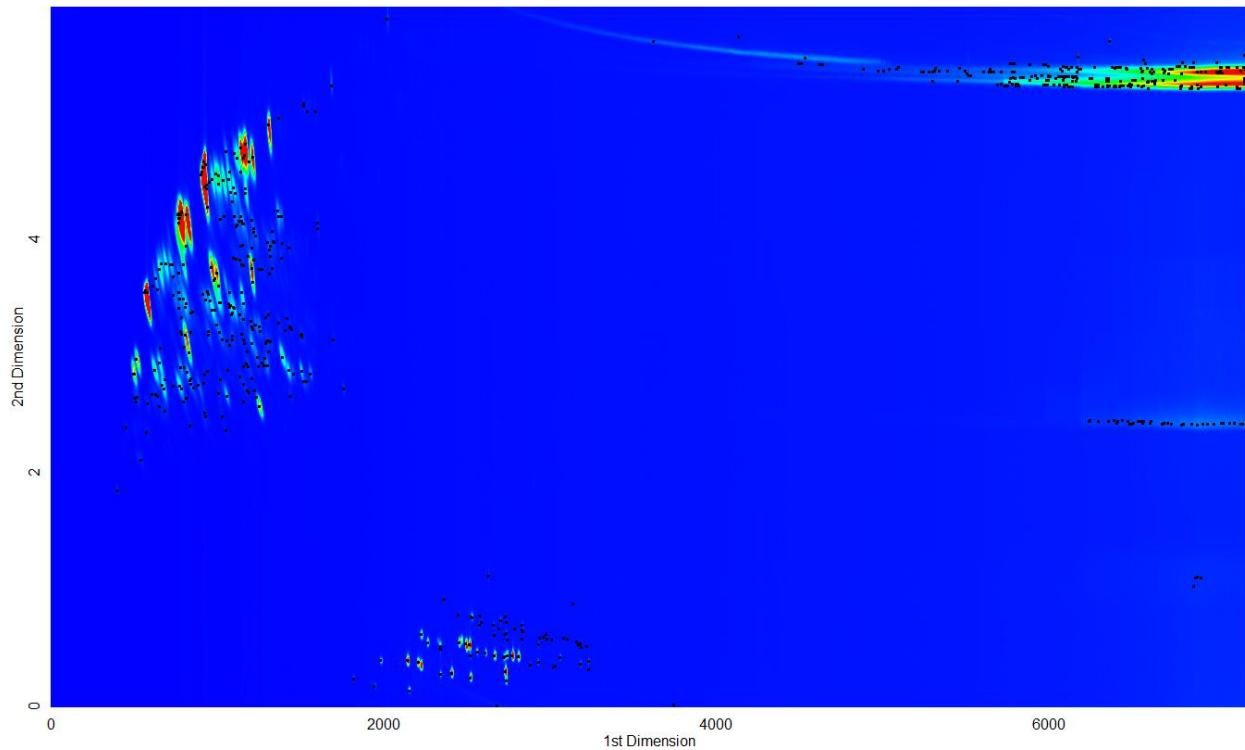
Column set 4: Sample VI



Column set 4: Sample VII



Column set 4: Sample VIII



Column set 4: Sample IX

