© Copyright 2018

Christopher E. Freye

Development of Instrumental and Chemometric Techniques for the Analysis of Complex Samples via Multi-Dimensional Gas Chromatography

Christopher E. Freye

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Robert Synovec, Chair

Matthew Bush

D. Michael Heinekey

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Development of Instrumental and Chemometric Techniques for the Analysis of Complex Samples via Multi-Dimensional Gas Chromatography

Christopher E. Freye

Chair of the Supervisory Committee:
Professor Robert E. Synovec
Department of Chemistry

A combination of four instrumental systems and five chemometric methods are shown to improve the efficiency and resolving power (i.e. peak capacity/ peak capacity production) of multi-dimensional gas chromatography (MDGC) as well as provide straightforward, easily interpretable chemical information. Implementation of a high speed pulse flow valve for two-dimensional gas chromatography (GC × GC) is shown to provide ultra-fast modulation with modulation periods (P_M) as short as 50 ms. Using a commercially available pulse flow valve, this injection technique performs a combination of vacancy chromatography and frontal analysis, whereby each pulse disturbance in the analyte concentration profile as it exits the first column (^{1}D) results in data that is readily converted into a second separation (^{2}D). A three-step process converts the raw data into

a format analogous to a GC × GC separation, incorporating signal differentiation, baseline correction and conversion to a GC × GC chromatogram representation. For a P_M of 250 ms, the apparent peak width on the ^{2}D, ^2W_b, ranged from 12 to 45 ms producing a ^{2}D peak capacity, ^2n_c, of ~ 10, and the total peak capacity, $n_{c,2D}$, was 4300 or a peak capacity production of 650 peaks/min. Next, the use of a high temperature diaphragm valve as a modulator for GC × GC facilitated separation temperatures up to 325 °C. Previous diaphragm valve technology limited use to 175 °C if the valve was mounted in the oven or to 265 °C if the valve was face mounted on the outside of the oven. A 44-component mixture was evaluated and the diaphragm valve created narrow, reproducible peaks on the ^{2}D dimension leading to a peak capacity production of 300 peaks/min and had minimal retention time shifting on the ^{1}D and ^{2}D dimensions. In addition, the high temperature diaphragm valve was shown to increase the detection sensitivity by ~8 times compared to one-dimensional gas chromatography due to zone compression. Furthermore, the high temperature diaphragm valve was proven to be compatible with time-of-flight mass spectrometry (TOFMS). Finally a combination of the high temperature diaphragm valve and pulse flow valve yielded a three-dimensional gas chromatography system (GC^3) that had a peak capacity production of 1000 peaks/min which is a ~5 times increase in efficiency compared to other GC^3 systems.

Investigation of novel chemometric techniques for the analysis of GC × GC is shown to be beneficial in extracting useful chemical information from complicated samples. Kerosene-based rocket fuels were analyzed via a GC × GC – FID system that implemented a high temperature diaphragm valve as a modulator. Using leave-one-out cross validation (LOOCV), the summed GC × GC – FID signal of three compound-class selective 2D regions (alkanes, cycloalkanes, and aromatics) was regressed against previously measured ASTM derived values. Additionally, a more

detailed partial least squares (PLS) analysis was performed on compound classes (*n*-alkanes, *iso*-alkanes, mono-, di-, and tri-cycloalkanes, and aromatics) as well as the physical properties previously determined by ASTM methods (such as net heat of combustion, hydrogen content, density, kinematic viscosity, sustained boiling temperature and vapor rise temperature). The resulting models had low root mean square errors of cross validation (RMSECV) and had similar outcomes to previously reported results using a GC × GC – TOFMS. Using the information gained from the study, a more extensive study of predicting four physical properties (e.g., viscosity, heat of combustion, hydrogen content, and density) was undertaken using 74 different kerosene-based fuels. Highly reliable PLS models were developed that related chemical composition obtained via GC × GC – TOFMS to fuel properties obtained via ASTM methods. The PLS prediction of the four physical properties (e.g., viscosity, heat of combustion, hydrogen content, and density) had relatively low errors of RMSECV values of 0.0434, 38.1, 0.112, and 0.0037, respectively. Investigation of the linear regression vectors (LRVs) indicate the relationship between the chemical composition and physical properties enabling the chemical compositions of fuels to be altered to meet certain industrial specifications. Using a similar fuel set comprised of 36 kerosene-based fuels, the thermal stability was evaluated using a novel instrument, CRAFTI (Compact Rapid Assessment of Fuel Thermal Integrity). Using a chemometrics-based feature discovery algorithim, the chemical information obtained via GC × GC – TOFMS was correlated to the thermal integrity data. Certain forms of carbon that were deposited within the test article of the CRAFTI instrument were found to strongly correlate with increased backpressure and some of the more prevalent compounds were identified. Next, tile-based Fisher ratio (F-ratio) analysis was applied to tandem ionization time-of-flight mass spectrometry (TI – TOFMS) in order to enhance discovery-based analyses. A hard ionization energy (70 eV) and soft ionization energy (14 eV)

were collected concurrently, and the discovery of 12 analytes spiked in diesel fuel was shown to be improved when the two ionization energies were used in tandem resulting in a higher discovery rate while also lowering the number of false positives. Using parallel factor analysis (PARAFAC) the analytes that were “discovered” were deconvoluted in order to obtain their identification via match values. Lastly, the limit of detection (LOD) and limit of quantification (LOQ) were improved by a novel integration method. Signal to noise (*S/N*) enhancement was theoretically studied using simulations and both the LOD and LOQ can be lowered by a factor of 3. When compared to the two-step, a commonly applied method for quantifying one-dimensional and two-dimensional, the integration method resulted in more accurate and precise measurements at low *S/N*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The journey through graduate school has been quite a tumultuous experience with plenty of highs and lows. I have to thank my graduate research advisor, Dr. Rob Synovec. Without your guidance and support, I would not have made it through graduate school. Secondly, thank you to Dr. Michael Sepaniak at the University of Tennessee at Knoxville who allowed me to perform undergraduate research. Performing research under your supervision prepared me for graduate school in many untold ways.

Thank you to the many colleagues I have had the privilege of working with over the past four years. A very large thank you to Dr. Brian Fitz who helped me set up my first GC × GC instrument and answered all my questions as I was first getting started. To Dr. Brendon Parsons who was willing to help me when I had questions about coding and introduced me to the perfect method of brewing coffee. To Dave Pinkerton for sitting down and discussing the nuances of my work, thank you for your advice and thoughtful suggestions. To Dan Baha, thank you for being my friend. Not only were we able to share knowledge and projects in graduate school, but our friendship outside of work is something I will miss. To Sarah Prebihalo and Kelsey Berrier, thank you for the many doughnut breaks, letting me rant, and the many ideas that furthered my research. Thank you to Dr. Brooke Reaser, Dr. Nate Watson, Nick Moore, Paige Sudol, and Derrick Gough for being great group members and scientific collaborators.

Finally, thank you to all my family and friends outside the University of Washington who over the years offered support and encouragement. There are so many of you who helped me on my journey. Thank you to my parents, who sacrificed so much so that I could succeed. Thank you to my sister, Jackie, for being my friend and always being there for me.

Chapter 1. Introduction to Gas Chromatography, Multi-Dimensional Gas Chromatography, and Chemometric Analyses

Some parts of this chapter have been reproduced from S.E. Prebihalo, K.L. Berrier, C.E. Freye, H.D. Bahaghighat, N.R. Moore, D.K. Pinkerton, and R.E. Synovec, "Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications" *Analytical Chemistry* 90 (2018) 505-532.

1.1 INTRODUCTION TO CHROMATOGRAPHY

1.1.1 *History and Overview of Chromatography*

Chromatography is an analytical technique that was invented at the turn of the 20th century by Mikhail Tswett, an Italian-born scientist studying plant pigments [1]. Tswett used a chalk-filled tube and organic solvent to physically separate chlorophyll, xanthophylls, and carotene (colored pigments). In his landmark paper in 1906, Tswett called his separation technique "chromatography" which in Greek, *chroma* means "color" and *graphein* means "to write." The evolution from the original separation technique to modern high performance separations occurred over the beginning of the 20th century, and in 1952 Martin and James's article on gas chromatography can be considered the beginning of modern, high-performance chromatography [2]. For a complete history of chromatography, the readers are referred to several publications [3-5]. There are many different forms of chromatography (e.g. liquid, gas, supercritical fluid, size-exclusion), but each has the overarching goal to physically separate a mixture into its separate components.

Gas chromatography (GC) is a separation technique that uses an inert gas, usually helium or hydrogen but sometimes nitrogen, as a mobile phase [6]. The gaseous mobile phase flows through a column with a liquid or polymer stationary phase and separates volatile and semi-volatile species based on their affinity for the stationary phase. There is a plethora of stationary phases available that mainly separate species based on their boiling point and/or polarity; however, there are some application specific columns which may be beneficial for certain mixtures. GC is capable of resolving thousands of compounds in a single analysis, and when coupled to an appropriate detector, can detect them at very low concentrations which makes GC useful in many disciplines. A GC separation can be monitored by any number of detectors, both universal and specific to sample type. One of the most commonly employed GC detectors is the flame ionization detector (FID). A FID operates by measuring the number of ions formed during the combustion of the compounds in a hydrogen flame where the number of ions is proportional to the concentration of the organic species in the gas stream. The other most commonly employed GC detector is an MS detector which may come in many different forms. All MS detectors share a similar mechanism: analytes entering the MS are ionized, fragmented and separated based on their mass-to-charge ratio (*m/z*). Analytes exhibit unique fragmentation patterns and mass spectra which can be used to identify the unknown molecules.

1.1.2 *GC Theory and Separation Fundamentals*

GC Theory

Some of the fundamental aspects of chromatography will be provided in order to sufficiently quantify the improvements made in this dissertation. For a more in-depth coverage of GC theory, a review by Reid and Synovec [7] covers sufficient GC theory. Several figures of merit such as resolution, peak capacity, and peak capacity production will be discussed in depth. For

further understanding of these and other principles of chromatographic figures of merit, the reader is directed to these resources [6,8-10].

As previously mentioned, compounds present in a mixture are separated based on their affinity for the stationary phase. After a sample has been introduced to the system, the components will equilibrate between the stationary phase and the mobile phase as the mobile phase flows through the column. The distribution coefficient, K_D, is defined as

$$K_D = \frac{[Analyte]_{SP}}{[Analyte]_{MP}} \quad (1.1)$$

where $[Analyte]_{SP}$ is the concentration of the analyte in the stationary phase and the $[Analyte]_{MP}$ is the concentration of the analyte in the mobile phase. A higher concentration of the analyte in the stationary phase relative to the mobile phase will result in a higher distribution coefficient, which in turn will result in more retention on the column. The K_D can be tuned by differences in stationary phase composition as well as temperature.

The retention factor, *k*, is a means of measuring the retention of an analyte on a chromatographic column. The retention factor is the ratio of the time the analyte spent in the stationary phase over the time the analyte spent in the mobile phase.

The retention factor, *k*, is defined as

$$k = \frac{moles\ Analyte_{SP}}{moles\ Analyte_{MP}} \quad (1.2)$$

which also can be written as

$$k = \frac{t_R - t_0}{t_0} \quad (1.3)$$

where t_R is the retention time, the amount of time an analyte spent on column, and t_0 is the dead time, the time it takes an unretained analyte to travel through the separation column. Analytes that spend more time in the stationary phase will have a higher *k* than analytes that spend less time in

the stationary phase. The distribution coefficient, K_D, and retention factor, k, are related through the factor β which is the ratio of the volume of the mobile phase, V_G, to the volume of the stationary phase, V_S, and is represented by the following equation

$$\beta = \frac{V_G}{V_S} \tag{1.4}$$

Thus the relationship between, K_D, k, and β is defined as

$$K_D = \beta k \tag{1.5}$$

As analytes elute from the column and are observed by the detector, they exhibit a Gaussian profile which is defined as

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{1.6}$$

Here, σ is the standard deviation, μ is the mean, and the fraction prior to the exponential is equal to the area and is often replaced by a constant A. The retention time of an analyte, t_R, is equal to μ and is the time at which the maximum response of an analyte is measured. The width of the peak can be measured at half height, $W_{1/2}$, or at baseline, W_b, where the peak width at baseline is defined as 4σ or ± 2σ from t_R. Figure 1.1 shows the relationship between all these parameters as well as the fraction of the total peak height or signal that occurs at various distances from t_R.

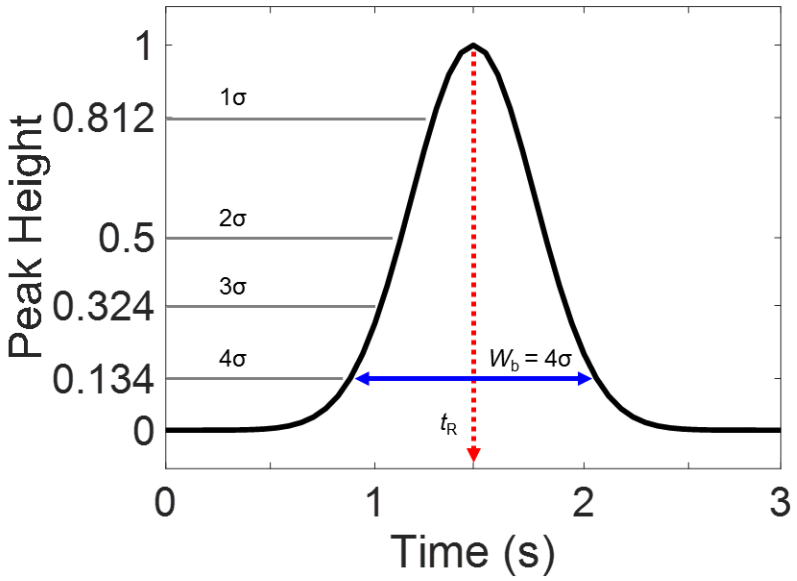


Figure 1.1 Gaussian peak with width at base, fractional peak heights and retention time identified.

GC Figures of Merit

A brief overview of GC theory has been provided, yet none of these parameters previously introduced define how efficient or effective a GC separation is. One way to calculate the separation between analytes is the resolution, R_s, defined as

$$R_s = \frac{t_{R,2} - t_{R,1}}{W_{b,avg}} \qquad (1.7)$$

Which is the difference in retention time between the two analytes divided by the average width at base of the two analytes as seen in Figure 1.2. A R_s of 1.5 is considered baseline resolution. A R_s of 1 is considered enough separation for quantitative analysis as there is only a 5% overlap. At a R_s of 0.5, the two analytes present will appear to a single peak, and at a R_s of 0.3, most deconvolution techniques begin to fail. Deconvolution will be discussed more in Section 1.2.4.

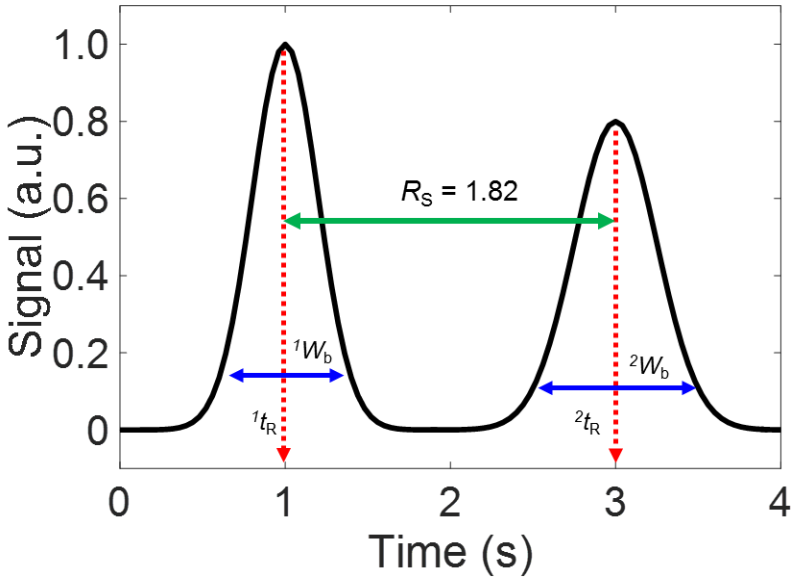


Figure 1.2 Visual representation of resolution between two peaks.

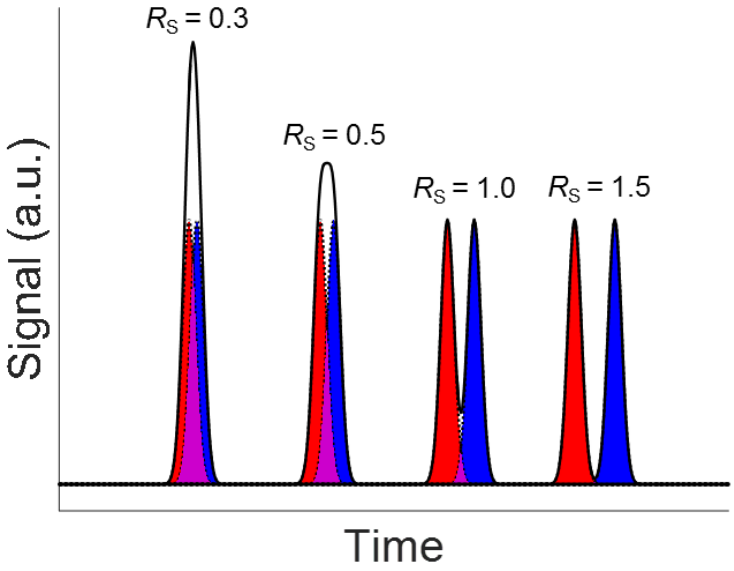


Figure 1.3 Representative peaks at various resolutions.

These four resolution values are shown in Figure 1.3 where the red and blue curves represent a single analyte, the thick black curve is what would be viewed by the detector response, and the purple area is the peak overlap due to the coelution. Indeed a chromatographer's goal is to maximum the resolution of a separation by tweaking various instrumental parameters or through use of chemometric techniques. Another method to improve resolution is to perform multi-dimensional separations which will be discussed in the next section.

In order to determine the efficiency or separation power of an instrumental setup, several metrics have been introduced. *N*, the number of theoretical plates, is often used to compare different columns with similar experimental setups, and plate height, *H*, has been extensively used in predicting peak widths from on-column band broadening using the Golay equation. The most commonly applied metric is peak capacity, n_c. Peak capacity is defined as the number of peaks with an average width at base (4σ), $W_{b,avg}$, that can be separated in a given separation window, $t_{sep.}$

$$n_c = \frac{t_{sep}}{W_{b,avg}} \qquad (1.8)$$

Peak capacity is a useful metric because it is easy to calculate from experimental data. By measuring a few peaks from the separation, a good understanding of the peak capacity of the instrumental system can be obtained. Normally, n_c, is calculated at unit resolution ($R_s = 1$). For a typical GC separation, peaks are ~ 3 seconds wide. With a 20 minutes (1200 seconds) separation window, the peak capacity is 400. Peak capacity can be improved in either two ways: 1) increase separation time or 2) decrease peak widths. However, the latter is preferred because this does not increase the duty cycle (e.g. total run time plus cool down time) of the instrument which could lead to a long analysis time when multiple samples are analyzed. Perhaps the most suitable method for determining the separation efficiency of an instrument is peak capacity production, which is peak capacity per unit time, defined as

$$\frac{n_c}{t_{sep}} = \frac{1}{W_{b,avg}} \tag{1.9}$$

Using the same numbers as the peak capacity example above (n_c = 400, t_{sep} = 20 min), the peak capacity production is 20 peaks/min.

1.1.3 *Multi-Dimensional Gas Chromatography (GC × GC and GC^3)*

Perhaps the most powerful way to increase peak capacity and peak capacity production is through the use of two- and three-dimensional separations. Giddings introduced the idea of comprehensive two-dimensional separations in 1984 [9], but it was not until 1991 when the first two-dimensional gas chromatograph (GC × GC) was developed by Liu and Phillips [11]. In GC × GC, two separation columns are employed with sufficiently orthogonal stationary phases (different chemical selectivity) and are connected through a modulation interface which transfers eluate from the first column to the second column with a user defined interval, termed the modulation period, P_M. Chapters 1, 2, and 3 explore the use of new modulators for GC × GC. The peak capacity of a GC × GC instrument is the combination of the peak capacities on each dimensions

$$^1n_c * \ ^2n_c = n_{c,2D} \tag{1.10}$$

Or more detailed

$$\frac{^1t_{sep}}{^1w_{b,avg}} * \frac{^2t_{sep}}{^2w_{b,avg}} = n_{c,2D} \tag{1.11}$$

where $^2t_{sep}$ is equal to the P_M. Because the separation times of a GC × GC are similar to a GC instrument, the addition of a second column significantly improves the peak capacity and peak capacity production relative to a one dimensional (1D) GC. Moreover, the additional selectivity added by the orthogonal stationary phase of the second dimension aids in separating compounds that may co-elute on the first dimension. Another important parameter for multi-dimensional gas

chromatography (MDGC) is the sampling density, ρ_s, [12] also known as the modulation ratio, M_R, [13] defined as

$$\rho_s = \frac{W_b}{P_M} \tag{1.12}$$

where W_b is the width at base on the first dimension and P_M is the modulation period. An analyst strives to keep this number between ~2-4 as undersampling can lead to loss of quantitative precision and modulator induced band broadening while oversampling can lead to a loss in ${}^2n_c$ [12-16]. Using the peak width from the 1D-GC calculations (${}^1W_b$ = 3 sec, t_{sep} = 20 min), an appropriate P_M is 1.5 seconds. Peak widths on the second dimension, ${}^2W_b$, can range from 40-200 ms wide. Due to the pseudo-isothermal nature of the second dimension, peaks become wider directly proportional to their retention, ${}^2k$, on the second dimension. Normally a ${}^2n_c$ ranging from ~8-15 is obtained which matches the theoretical predictions by Blumberg and co-workers [17].

Comprehensive three-dimensional gas chromatography (GC^3) was pioneered by Watson et al in 2007 [18]. Improvements to instrumental design by using new state-of-the-art modulators and hyphenation to time-of-flight mass spectrometers (TOFMS) and data analysis techniques have been recently published [19,20]. When compared to GC × GC, GC^3 does not have the expected increase in peak capacity from adding a third dimension of separation. This is mainly due to the limitations of the modulators that were implemented and the data collection frequencies. While the peak capacities from GC^3 and GC × GC are similar, GC^3 offers an additional dimension of chemical selectivity [21]. Indeed, using a third column may allow the analyst to tune the separation in order to focus on critical regions of interest. The calculation of peak capacity for a GC^3 instrument is an extension of the peak capacity calculation for GC × GC

$$\frac{{}^1t_{sep}}{{}^1w_{b,avg}} * \frac{{}^2t_{sep}}{{}^2w_{b,avg}} * \frac{{}^3t_{sep}}{{}^3w_{b,avg}} = n_{c,3D} \tag{1.13}$$

Chapter 4 explores the combination of the work in Chapters 1 and 2 to create a novel GC^3 instrument which has a large peak capacity production.

1.2 DATA ANALYSIS AND CHEMOMETRICS

1.2.1 *Introduction to Chemometrics*

The ability to analyze GC, GC × GC, and GC^3 data and obtain useful chemical information is paramount to the utility of separations. All three, especially GC × GC and GC^3, produce large amounts of data, but when these separation techniques are coupled with multichannel detectors such as a TOFMS, the amount of data is enormous. Indeed, when multiple samples or replicates are analyzed, it can become a hassle to extract meaningful information. There are challenges in working with such complex and dense data sets. Some challenges include analyte peak overlap, difficult and time-consuming computations, confident discovery of important (and often subtle) chromatographic differences between samples, and extraction of other meaningful information. Optimization of the experimental and instrumental conditions is very important in dealing with these issues. However, when instrumental capabilities have been fully optimized and have become a limiting constraint, the analyst must apply other approaches to achieve the desired goals. The application of chemometrics can provide the analyst with a means to address these goals. The field of chemometrics specializes in mathematically gleaning useful information from chemical measurements [22-27]. However, specific care must be taken for some chemometric methods as some techniques have specific requirements such as bilinearity or trilinearity and suitable data density across the ^1D and ^2D peaks. Many of these chemometric methods require a minimum amount of chemical selectivity, in both the separation (i.e., adequate resolution) and detection dimensions. The common classes of chemometrics include property prediction, pattern

recognition, deconvolution, and retention time/index modeling. A brief overview of property prediction, pattern recognition, deconvolution is provided below.

1.2.2 *Property Prediction*

There is growing need to relate physical and/or chemical property data obtained via a variety of other methods (some analytical, some not so much) to the chemical information obtained from MDGC. Property prediction is also important to link the chemical information contained in the chromatographic data to the property data, in order to create meaningful understandings between the two data sets.

Partial Least Squares (PLS)

Partial least-squares (PLS) regression analysis is the most commonly implemented chemometric method for the prediction of properties. A detailed review of the theory of PLS analysis can be found elsewhere [28], but briefly, PLS analysis attempts to correlate two data matrices (X-block and Y-block) by calculating loadings referred to as the number of latent variables (LVs). PLS analysis aims to model the covariance between these two matrices by finding the multidimensional direction in the X-block which explains the maximum multidimensional variance in the Y-block. In other words, PLS analysis uses analytes which have a large range of intensities across the MDGC samples and attempts to correlate them to the differences in the physical or chemical properties of the samples (the property data). PLS modeling provides a one-to-one (i.e. linear) correspondence between the measured property values relative to the predicted property values. While the model is a linear correlation, it is possible to scale the data (e.g. logarithmically, quadratically, etc.). PLS analysis provides the following two valuable outcomes. First, using a training set of samples, a linear correspondence of the previously measured property data is modeled using the MDGC data, so subsequent analyses of MDGC data of new samples can

be used to predict the property value of the new samples without having to directly measure these properties on the new samples. Second, since the underlying relationship between the chemical compositions of the samples is correlated to the property data, a deeper understanding between the two measurement platforms is provided. The first outcome of the PLS modeling is the regression plot showing the relationship between the MDGC dataset and the predicted property. This is typically shown as the measured property on the *x*-axis and predicted property on the *y*-axis. The second outcome of the PLS modeling is the linear regressions vectors (LRVs), which can be examined to determine which compounds are positively correlated (or negatively correlated) with the property being predicted. Analytes that exhibit positive values in the LRVs are correlated with increasing the specified property while negatively correlated analytes will exhibit negative values in the LRV and are correlated with decreasing the specified property.

1.2.3 *Pattern Recognition*

A major goal in performing MDGC analyses is to discover relevant features in the data across multiple samples. Many pattern recognition applications fall under this category of cross-sample analysis, including sample classification, class comparison, chemical fingerprinting, and chemical/biomarker discovery. These methods are classified as nontargeted approaches, where the relevant analytes may be unknown beforehand, and the experimental design may be either supervised or unsupervised. Supervised methods are used when information about the samples (i.e., classes) is known beforehand and thus utilized in the experimental design and operation of the method. Unsupervised methods are applied when there is much less known about the sample classifications *a priori* and this information is desired (i.e., sample class membership). Classification is used to describe pattern recognition methods that aim to classify samples into groups based on similarity such as principal component analysis (PCA). Class comparison

methods aim to identify significant differences between samples belonging to different classes, whereby class membership is known *a priori*, such as Fisher ratio (F-ratio) analysis, which can assist/result in the discovery of biomarkers or chemical markers that differentiate sample classes. Pattern recognition methods can be used for feature selection or data reduction, followed by further chemometric analysis.

Principal Component Analysis (PCA)

Principal Components Analysis (PCA) is an unsupervised nontargeted classification method that can be used to quickly analyze and compare a large number of samples and objectively extract useful information. More detailed information about PCA can be found elsewhere [29-31,32]. Briefly, PCA translates the information from an n-dimensional variable into principal component (PC) space where there is one score describing each variable in the data set in the new, much lower dimensional, PC space. Mathematically, PCA is described as:

$$X = TDP^T + E \qquad (1.14)$$

where X is the data matrix with m rows and n columns, with m equal to the number of samples and n equal to the number of variables. T is the matrix whose columns consist of the scores vectors, t_i, and P^T is the matrix whose rows are loadings vectors, p^T_i. D is the diagonal matrix with the diagonal elements equal to the square roots of the eigenvalues of X^TX, and E is the matrix of residuals. The PCs are sorted based on their eigenvalues and oriented such that they span the greatest variation inherent in the scores. Thus the eigenvalue quantifies the amount of variance (i.e. chemical/physical information) captured by the PC. PCs that contain useful chemical/physical information are kept while PCs that contain predominately noise are discarded. PCA results in a scores plot which shows the degree of sample clustering, shedding information on distinct grouping of samples and loadings plots which provide information about the basis of the

groupings; in other words, the loadings can be used to determine which regions of the chromatogram (i.e., peaks/analytes) are important in differentiating the various sample classes.

Fisher Ratio Analysis (F-ratio)

Fisher ratio (F-ratio) analysis, initially described by Fisher [33], is a supervised nontargeted approach to discover underlying differences in samples. Supervision refers to prior classification of samples/chromatograms as they relate to the experimental design. The F-ratio approach is an analysis of variance (ANOVA) method that provides data reduction to elucidate class distinguishing features. The F-ratio is defined as:

$$F - ratio = \frac{\sigma_{bc}^2}{\sigma_{wc}^2} \qquad (1.15)$$

where σ_{bc}^2 is the variance present between the classes and σ_{wc}^2 is the variance within the classes. Since the F-ratio calculation is based on the variance of the signals in the MDGC data (raw data approach) or variance in quantified analytes (peak table approach), the F-ratio method prioritizes statistical significance over absolute signal/concentration with a higher F-ratio having a greater difference between the two classes.

An analyst can apply the F-ratio method to MDGC data in three different ways: pixel-based, peak table-based, and tile-based. Pixel-based F-ratio analysis means an F-ratio is calculated at every data point within the chromatogram, i.e., at every point in which a mass spectrum has been collected in the multi-dimensional separation [34,35]. However, pixel-based F-ratio may suffer from false positives or false negatives due to the misalignment of the data. This can be mitigated by aligning the data prior to analysis. For peak table-based F-ratio, peak tables are calculated, the tables are aligned, and then the F-ratio is calculated [36]. While this is a simpler method then dealing with 2D retention time misalignment, peak tables may miss analytes present at low concentrations. Another way to deal with misalignment is binning the data based upon

the ^{1}D and ^{2}D peak widths in relation to the observed retention time shifting. However, this will result in splitting analytes into multiple bins and lower F-ratio values. In order to address this issue with binning, a method has been introduced to use a tiling scheme which creates bins at different locations enabling the bin to be centered on the peak resulting in a maximum F-ratio [37-40]. Chapter 8 will explore using the tile-based F-ratio method in conjunction with a new type of detector for improved discovery-based analysis. After the F-ratios are calculated, a hit list is assembled with the highest F-ratio indicative of the statistically most prominent feature or "hit". In order to determine a point at which the analysis becomes unreliable, a null distribution approach which takes advantage of the inherent noise (instrumental and chemical) in the data set will be implemented [38-40].

1.2.4 *Deconvolution Techniques*

Despite the increased peak capacity due to additional separation dimension(s), analyte overlap is ubiquitous in GC × GC (or GC^{3}) separations. Chromatographic peak deconvolution methods play an integral role in addressing this challenge to provide confident analyte identification, quantification, and peak purity assessment. Through the application of deconvolution methods, such as parallel factor analysis (PARAFAC), overlapped peaks can be computationally separated on the ^{1}D and ^{2}D (and ^{3}D) time and spectral dimensions in order to gain chemical information for specific analytes. With the deconvoluted peak data, analytes that were overlapped can be identified using the deconvoluted spectra and quantified with the peak profiles obtained, thereby facilitating calibration methods [41].

Parallel Factor Analysis (PARAFAC)

Parallel factor analysis (PARAFAC) is a tensor rank decomposition method that can be applied to three-way (or greater) data (e.g. GC × GC – TOFMS) in order to decompose the data

into components that ideally describe the number of analytes present [42]. PARAFAC mathematically resolves a given trilinear structure by modeling each component as the outer product of its three fixed vectors (i.e. ${}^{1}t_{R}$, ${}^{2}t_{R}$, *m/z*), while also simultaneously filtering out nontrilinear signals such as noise contributions to signals. Mathematically PARAFAC is:

$$\underline{R} = \sum_{for\ i=1\ to\ n} x_i \times y_i \times z_i + \underline{E} \qquad (1.16)$$

where $\underline{R}$ represents the data cube generated by GC × GC – TOFMS separation. Each of the n components is resolved into the x, y, and z vectors which represent the first dimension profile, second dimension profile, and mass spectral profiles, respectively. Any of the remaining signal is retained in the residual matrix, $\underline{E}$. PARAFAC has a strict trilinearity requirement which means as the data diverges further from this constraint, the PARAFAC model will diverge from the "real" answer [43,44]. More detailed about PARAFAC can be found in these suggested references [45,46].

1.3 CHALLENGES AND MOTIVATIONS

In order to place this dissertation in the context of the field of MDGC, it is important to address some of the challenges within the field and the motivations to further investigate certain areas. The goal of MDGC, as with any analytical technique, is to provide the greatest amount of information within the shortest period of time. This prerequisite can be met in two different approaches, or more likely a combination of both. The first approach is to use a state-of-the-art instrument resulting in high peak capacity production and high chemical selectivity in conjunction with a selective detector such as a TOFMS. The second approach is to use chemometrics in order to extract useful and meaning chemical information from the data rich datasets. Chapters 2, 3, and 4 introduce new modulators, pulse flow valves and high temperature diaphragm valves, for GC × GC in an attempt to improve peak capacity production. Chapter 5 uses the information gained in

Chapters 2-4 and applies it GC^3 instrumentation in order to create a highly chemically selective instrument with the added benefit of high peak capacity production. Chapter 6 is a combination of state-of-the-art instrumentation and chemometrics in order to correlate the physical properties of kerosene-based rocket fuel to the chemical information obtained via GC × GC. Using the knowledge gained in Chapter 6, Chapter 7 expands upon the study using a large dataset of 74 different aerospace grade fuels to gain insight into the relationship between chemical composition of the fuels and their physical properties. In Chapter 8, the chemical information obtained from GC × GC – TOFMS is correlated to the physical properties obtained from CRAFTI, a simulative environment for regenerative cooling, in order to gain an understanding of the relationship between thermal stability of a fuel and its chemical composition. Chapter 9 investigates the complementary information provided by tandem ionization time-of-flight-mass spectrometry (TI – TOFMS) in order to enhance discovery-based analysis. Furthermore, it is demonstrated how an analyst can implement a work flow creating a streamlined process from experimental to design to extraction of meaningful information. Lastly, Chapter 10 provides a new strategy for improving the limits of detection (LOD) and limits of quantification (LOQ) and compares this new strategy to the established two-step technique.

1.4 HYPOTHESES

The following chapters describe the investigative research performed over the past five years, which focus on new modulators for two-dimensional and three-dimensional gas chromatography and various chemometric methods for analyzing complex datasets. A brief summary of each chapter is provided below.

1.4.1 *Chapter 2: Comprehensive Two-Dimensional Gas Chromatography Using Partial Modulation via a Pulsed Flow Valve with a Short Modulation Period*

Partial modulation via a pulsed flow valve for comprehensive two-dimensional (2D) gas chromatography (GC × GC) is demonstrated, producing narrow peak widths, ${}^{2}W_{b}$, on the secondary separation dimension, ^{2}D, coupled with short modulation periods, P_M, thus producing a high peak capacity on the ^{2}D dimension, ${}^{2}n_c$. The GC × GC modulator is a pulse flow valve that injects a pulse of carrier gas at the specified P_M, at the connection between the primary, ^{1}D, column and the ^{2}D column. Using a commercially available pulse flow valve, this injection technique performs a combination of vacancy chromatography and frontal analysis, whereby each pulse disturbance in the analyte concentration profile as it exits the ^{1}D column results in data that is readily converted into a ^{2}D separation. A three-step process converts the raw data into a format analogous to a GC × GC separation, incorporating signal differentiation, baseline correction and conversion to a GC × GC chromatogram representation. A 115-component test mixture with a wide range of boiling points (36–372 °C) of nine compound classes is demonstrated using modulation periods of P_M = 50, 100, 250, and 500 ms, respectively. For the test mixture with a P_M of 250 ms, peak shapes on ^{2}D are symmetric with apparent ${}^{2}W_{b}$ ranging from 12 to 45 ms producing a ${}^{2}n_c$ of ~ 10. Based on the average peak width of 0.93 s on the ^{1}D separation for a time window of 400 s, the ^{1}D peak capacity is ${}^{1}n_c \sim 430$. Thus, the ideal 2D peak capacity $n_{c,2D}$ is 4300 or a peak capacity production of 650 peaks/min using the P_M of 250 ms. Additionally, for a P_M of 50, 100 and 500 ms, the ${}^{2}n_c$ are 4, 7, and 12, respectively. Retention times on ^{2}D, ${}^{2}t_R$, are reproducible having standard deviations less than 1 ms. Finally, the processed data is shown to be quantitative, with an average RSD of 4.7% for test analytes.

1.4.2 *Chapter 3: High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography*

A high-temperature diaphragm valve-based comprehensive two-dimensional gas chromatography (GC × GC) instrument is demonstrated which readily allows separations up to 325 °C. Previously, diaphragm valve-based GC × GC was limited to 175 °C if the valve was mounted in the oven, or limited to 265 °C if the valve was faced mounted on the outside of the oven. A new diaphragm valve has been commercially developed, in which the temperature sensitive O-rings that previously limited the separation temperatures have been replaced with Kalrez O-rings, a perfluoroelastomer, allowing for significantly higher temperatures permitting a greater range of volatile and semi-volatile compounds to be readily separated. In the current investigation, a separation temperature up to 325 °C is demonstrated with the valve mounted directly in the oven. Since the temperature limit for most commonly used GC columns is at or below 325 °C, the scope of diaphragm valve-based GC × GC is now dramatically broadened to encompass a majority of all column stationary phase chemistries. A 44-component mixture of alkanes, alcohols, and polyaromatic hydrocarbons is used to study this new configuration whose boiling points range from 98 °C (n-heptane) to 450 °C (n-triacontane). For the test mixture using a modulation period P_M of 1.0 s, peak shapes on second dimension separations, ^{2}D, are symmetric with average widths at base of 79.4 ms, producing a ^{2}D peak capacity of $^2n_c \sim 12$. Based on the average peak width of 2.4 s for the first dimension separation with a run time of 32.5 min, the ^{1}D peak capacity is $^1n_c \sim 800$. Thus, the ideal two-dimensional peak capacity $n_{c,2D}$ is 9600. Little variation in within-analyte ^{2}D peak width was observed with an average %RSD of less than 3.0%. Furthermore, retention time on ^{2}D was very reproducible with an average %RSD less than 0.5%. Measured peak areas (sum of all ^{2}D peaks for given analyte) had an average %RSD of 4.4%. The

transfer fraction from ^{1}D to ^{2}D was experimentally determined to be ~30%, while the detection sensitivity for valve-based GC × GC was ~8 times higher than one dimensional GC due to zone compression. After a year of use with temperatures consistently up to 325 °C, there has been no deterioration of the valve or its performance for GC × GC. Separations of vacuum pump oil and orange oil are also reported to demonstrate practical utility.

1.4.3 *Chapter 4: High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry*

A high temperature diaphragm valve-based two-dimensional (2D) gas chromatography (GC × GC) with time-of-flight mass spectrometry (TOFMS) instrument, with the valve mounted directly in the GC oven, is demonstrated with separations up to 325 °C. Use of the diaphragm valve allowed for the use of uncoupled carrier gas flows for ^{1}D (first column dimension) and ^{2}D (second column dimension), with a ^{1}D flow rate of 1 ml/min, and a ^{2}D flow rate 3 ml/min. The 3 ml/min flow rate on ^{2}D was selected to ensure compatibility with most TOFMS detectors. For valve-based modulation, a 5 µl sample loop coupled with a 60 ms pulse width was selected, providing sufficient sensitivity concurrent with an acceptable ^{2}D peak capacity. A 44-component mixture of alkanes, alcohols, and polyaromatic hydrocarbons (including n-alkanes of heptane to triacontane) whose boiling points range from 98 °C to 450 °C was used to initially study instrument performance. For a 120 min separation and a modulation period P_M of 2 s, average peak widths-at-base of 10 s and 94 ms were achieved for the alkanes on the ^{1}D and ^{2}D dimensions, respectively. Hence, the ^{1}D peak capacity is $^{1}n_c$ ~ 700, and the ^{2}D peak capacity can in principle be $^{2}n_c$ ~ 20. Thus the ideal 2D peak capacity could in principle approach $n_{c,2D}$ ~ 14,000 using the 120 min ^{1}D run time. The limit of detection (LOD) for docosane, a representative analyte, was determined to be 1.5 ppm injected concentration when 2 µl of liquid sample was injected with a

20:1 split. A separation of diesel fuel demonstrated the practical utility of the instrument with this complex sample using a relatively fast run time of 20 min and a short modulation period P_M of 1 s. Average peak widths-at-base of 5.4 s and 166 ms were achieved on the ^{1}D and ^{2}D dimensions, respectively. This yielded a $^1n_c \sim 220$ and a $^2n_c \sim 6$. Therefore, the 2D peak capacity is $n_{c,2D} \sim 1320$ with the 20 min diesel separation.

1.4.4 *Chapter 5: Investigation of Ultrafast Separations via a Pulse Flow Valve for MDGC (GC × GC, GC^3)*

Ultrafast modulation (P_M < 75 ms) via a pulse flow valve is demonstrated for comprehensive two-dimensional gas chromatography (GC × GC) and comprehensive three-dimensional gas chromatography (GC^3). Improving on previous work, the separation conditions are optimized for the pulse flow valve resulting in significant increases in peak capacity and peak capacity production. Using GC × GC, a 115-component test mixture is evaluated using a P_M of 500 creating average ^2W_b of 25 ms producing a ^2n_c of ~ 20. Based on the average peak width of 1.0 s on first (^{1}D) dimension, the total peak, $n_{c,2D}$, is ~7200 or a peak capacity production of ~1200 peaks/min was achieved. Using the optimized conditions, a P_M of 75 ms is evaluated producing ^2W_b of 8 ms. Using the knowledge gained from the GC × GC instrumental experiments, the pulse flow valve was implemented as the second modulator in a GC^3 instrument. A simple volatile mixture containing 18 compounds, the 115-component test mixture, and diesel fuel spiked with 8 compounds where analyzed by the GC^3 instrument. Evaluating the GC^3 instrumental configuration with the 115-component test mixture with a ^1P_M of 1.2 s and a ^2P_M of 60 ms is demonstrated. The peak shapes on the ^{1}D, ^{2}D and ^{3}D are symmetrical with an average ^1W_b of 3.2 s, ^2W_b of 130 ms, and ^3W_b 13 ms producing a ^1n_c of ~ 220, ^2n_c of ~ 9.5, and ^3n_c of ~ 5. Experimentally a total peak capacity, $n_{c,3D}$, of ~ 10,600 is achieved for the 4 representative analytes for a 10.8 min separation,

which results in a peak production of 1000 peaks/min. The analytical benefits of employing three varied chemical selectivity of the 3D separation is illustrated by the detection of spiked analytes as part of diesel fuel analysis.

1.4.5 *Chapter 6: Partial Least Squares Analysis of Rocket Propulsion Fuel Data Using Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography Coupled with Flame Ionization Detection*

The chemical composition and several physical properties of RP-1 fuels were studied using comprehensive two-dimensional (2D) gas chromatography (GC × GC) coupled with flame ionization detection (FID). A "reversed column" GC × GC configuration was implemented with a RTX-wax column on the first dimension (^{1}D), and a RTX-1 as the second dimension (^{2}D). Modulation was achieved using a high temperature diaphragm valve mounted directly in the oven. Using leave-one-out cross-validation (LOOCV), the summed GC × GC – FID signal of three compound-class selective 2D regions (alkanes, cycloalkanes, and aromatics) was regressed against previously measured ASTM derived values for these compound classes, yielding root mean square errors of cross validation (RMSECV) of 0.855, 0.734, and 0.530 mass%, respectively. For comparison, using partial least squares (PLS) analysis with LOOCV, the GC × GC – FID signal of the entire 2D separations was regressed against the same ASTM values, yielding a linear trend for the three compound classes (alkanes, cycloalkanes, and aromatics), yielding RMSECV values of 1.52, 2.76, and 0.945 mass%, respectively. Additionally, a more detailed PLS analysis was undertaken of the compounds classes (*n*-alkanes, *iso*-alkanes, mono-, di-, and tri-cycloalkanes, and aromatics), and of physical properties previously determined by ASTM methods (such as net heat of combustion, hydrogen content, density, kinematic viscosity, sustained boiling temperature and vapor rise temperature). Results from these PLS studies using the relatively simple to use and

inexpensive GC × GC – FID instrumental platform are compared to previously reported results using the GC × GC – TOFMS instrumental platform.

1.4.6 *Chapter 7: Gaining a Fundamental Understanding of Fuel Properties by Chemometric Analysis of Advanced Chemical Composition Measurement Data*

Kerosene-based jet and rocket fuels such as Jet A, JP-8, RP-1, and RP-2 are ubiqitous to the aerospace propulsion industry. There is an ongoing need to more fully assess the fuel composition (i.e., specific chemical compounds and compound classes present as a result of refining and feedstock blending) for the purpose of improving quantitative connections between fuel composition, properties, and ultimately system performance. Prediction of fuel performance through modeling puts a greater emphasis on obtaining detailed and accurate fuel property and compositional information. In order to optimally glean chemical composition information from complex fuels, comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry (GC × GC – TOFMS) has been shown to be an analytical effective tool. Additionally, multivariate data analysis methods, referred to as chemometrics, must be implemented in order to relate the vast, detailed chemical information captured within the GC × GC – TOFMS data, to complementary data obtained from the same fuels using ASTM standard measurement methods that infer fuel performance properties (e.g., viscosity, heat of combustion, hydrogen content, and density). In this study, a broad compositional range of distillate and multicomponent aerospace fuels is targeted, to gain a deeper understanding of the fuel physical properties. With an extremely large dataset of 74 different fuels, highly reliable partial least squares (PLS) models are developed enabling a comprehensive understanding between chemical composition obtained via GC × GC – TOFMS and fuel properties obtained via ASTM methods. The PLS prediction of the four physical properties (e.g., viscosity, heat of combustion, hydrogen content, and density) had relatively low

errors of RMSECV values of 0.0434, 38.1, 0.112, and 0.0037, respectively. This corresponds to an NRMSECV of 6.10%, 12.8%, 10.5%, and 7.81% for viscosity, heat of combustion, hydrogen content, and density, respectively. Investigation of the linear regression vectors (LRVs) indicate the relationship between the chemical composition and physical properties enabling the chemical compositions of fuels to in principle be altered to meet desired application specifications.

1.4.7 *Chapter 8: Recent Advances in the Development of Chemical Analysis Tools to Relate Compositional Variation to Thermal Integrity Data for RP-1, RP-2, and Related Fuels*

There is an ongoing need to more fully assess and control the composition of kerosene-based rocket propulsion fuels (i.e., specific chemical compounds and compound classes present as a result of feedstock blending and processing), in order to make better quantitative connections between fuel performance and fuel properties. Recent efforts in predicting fuel performance through modeling have put greater emphasis on detailed and accurate fuel properties and fuel compositional information. In order to optimally glean chemical composition information from these complex fuels, comprehensive two-dimensional gas chromatography combined with time-of-flight mass spectrometry (GC × GC – TOFMS) has been implemented. In this report, by applying multivariate data analysis techniques, referred to as chemometrics, the chemical compositional information from fuels provided using GC × GC – TOFMS is correlated to data obtained from the same fuels using CRAFTI (Compact Rapid Assessment of Fuel Thermal Integrity) which has been developed to simulate regenerative cooling channel environments in liquid rocket engines. Chemometric analysis of the data will be shown to provide significant insight into the thermal integrity of the fuels (manifested in the CRAFTI data: pressure and carbonaceous deposition profiles), and how the thermal integrity depends upon chemical composition (manifested in the GC × GC – TOFMS data). A recently developed chemometrics-

based chemical feature discovery approach correlates thermal integrity data to key chemical compounds in the GC × GC – TOFMS data, such as sulfur containing compounds, aromatic compounds, and other key chemical compounds. This research platform should have broad implications for the development of high fidelity composition-property models, leading to an optimized approach to fuel formulation and specification for advanced engine cycles.

1.4.8 *Chapter 9: Enhancing the Chemical Selectivity in Discovery-based Analysis with Tandem Ionization Time-of-Flight Mass Spectrometry Detection for Comprehensive Two-Dimensional Gas Chromatography*

The complementary information provided by tandem ionization time-of-flight mass spectrometry (TI – TOFMS) is investigated for comparative discovery-based analysis, when coupled with comprehensive two-dimensional gas chromatography (GC × GC). The TI conditions implemented were a hard ionization energy (70 eV) concurrently collected with a soft ionization energy (14 eV). Tile-based Fisher ratio (F-ratio) analysis is used to analyze diesel fuel spiked with twelve analytes at a nominal concentration of 50 ppm. F-ratio analysis is a supervised discovery-based technique that compares two different sample classes, in this case spiked and unspiked diesel, to reduce the complex GC × GC – TI – TOFMS data into a hit list of class distinguishing analyte features. Hit lists of the 70 eV and 14 eV data sets, and the single hit list produced when the two data sets are fused together, are all investigated. For the 70 eV hit list, eleven of the twelve analytes were found in the top thirteen hits. For the 14 eV hit list, nine of the twelve analytes were found in the top nine hits, with the other three analytes either not found or well down the hit list. As expected, the F-ratios per *m/z* used to calculate each average F-ratio per hit were generally smaller fragment ions for the 70 eV data set, while the larger fragment ions were emphasized in the 14 eV data set, supporting the notion that complementary information was provided. The

discovery rate was improved when F-ratio analysis was performed on the fused data sets resulted in eleven of the twelve analytes being at the top of the single hit list. Using PARAFAC, analytes that were "discovered" were deconvoluted in order to obtain their identification via match values (MV). Location of the analytes and the "F-ratio spectra" obtained from F-ratio analysis were used to guide the deconvolution. Eight of the twelve analytes where successfully deconvoluted and identified using the in-house library for the 70 eV data set. PARAFAC deconvolution of the two separate data sets provided increased confidence in identification of "discovered" analytes. Herein, we explore the limit of analyte discovery and limit of analyte identification, and demonstrate a general workflow for the investigation of key chemical features in complex samples.

1.4.9 *Chapter 10: Expanding the Limits of Detection and Quantitation in Chromatography through Integration*

The limit of detection (LOD) conventionally defined is three times the standard deviation of the noise while the limit of quantitation (LOQ) is ten times the standard deviation of the noise. These values are highly dependent on the noise and sensitivity of the detector as well as peak shape which can become wider with increased retention. Furthermore peak shape can become asymmetric by fronting or tailing due to column interactions resulting in a lower signal. By using this integration method, these limitations can be overcome providing better LOD and LOQ values that are independent of peak shape and is not hindered by Gaussian or low frequency noise. Through integration, the signal is transformed from univariate to a function dependent on time which incorporates the concentration profile and results in a more sensitive measurement. However, the noise has also been translated to a function of time and a new definition of noise has been presented for the integration method. Signal to noise (*S/N*) enhancement has been theoretically studied using simulations. Both LOD and LOQ enhancement are a function of peak

width (W_b), and both can be lowered by a factor 3 for a W_b of 20 data points which is commonly found in most separation methods. Herein, the two step method, commonly used for quantifying one dimensional and two-dimensional chromatography, is compared to the integration method. At high signal to noise (*S/N*), both techniques provide similar peaks areas while the integration method results in more accurate and precise measurements of peak area at low *S/N*. Furthermore, the integration method was studied for data that was collected via two dimensional gas chromatography with a time-of-flight mass spectrometer (GC × GC – TOFMS). The *S/N* improvement exceeded theoretical predictions and the integration method produced a 36-fold increase in *S/N*.

1.5 REFERENCES

[1] K.R. Williams, Colored Bands: History of Chromatography, J. Chem. Educ. 79 (2002) 922. doi:10.1021/ed079p922.

[2] A.T. James, A.J.P. Martin, Gas-liquid partition chromatography: the separation and micro-estimation of volatile fatty acids from formic acid to dodecanoic acid, Biochem J. 50 (1952) 679–690.

[3] V. Heines, Chromatography-history of a parallel development, Chem. Technol. 1 (1971) 280–285.

[4] L.S. Ettre, Development of chromatography, Anal. Chem. 43 (1971) 20A–31a. doi:10.1021/ac60308a022.

[5] G. Zweig, J. Sherma, Paper Chromatography—Past, Present and Future, J. Chromatogr. Sci. 11 (1973) 279–283. doi:10.1093/chromsci/11.6.279.

[6] R.L. Grob, E.F. Barry, eds., Modern Practice of Gas Chromatography, 4th ed., John Wiley & Sons, Inc., n.d.

[7] V.R. Reid, R.E. Synovec, High-speed gas chromatography: The importance of instrumentation optimization and the elimination of extra-column band broadening, Talanta 76 (2008) 703–717. doi:10.1016/j.talanta.2008.05.012.

[8] K. Robards, P.R. Haddad, P.E. Jackson, eds., Principles and Practice of Modern Chromatographic Methods, Academic Press, Boston, 2004, doi:10.1016/B978-0-08-057178-2.50001-7.

[9] J.C. Giddings, Unified separation science, Wiley, New York, 1991

[10] J.M. Miller, Chromatography, John Wiley & Sons, Inc., 2009

[11] Z. Liu, J. B. Phillips, Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, J. Chromatogr. Sci. 29 (1991) 227–231. doi:10.1093/chromsci/29.6.227.

[12] L.M. Blumberg, Accumulating resampling (modulation) in comprehensive two-dimensional capillary GC (GC×GC), J. Sep. Science 31 (2008) 3358–3365. doi:10.1002/jssc.200800424.
[13] W. Khummueng, J. Harynuk, P.J. Marriott, Modulation ratio in comprehensive two-dimensional gas chromatography, Anal. Chem. 78 (2006) 4578–4587. doi:10.1021/ac052270b.
[14] D.K. Pinkerton, B.A. Parsons, R.E. Synovec, Method to determine the true modulation ratio for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1476 (2016) 114–123. doi:10.1016/j.chroma.2016.11.015.
[15] J.M. Davis, D.R. Stoll, P.W. Carr, Effect of first-dimension undersampling on effective peak capacity in comprehensive two-dimensional separations, Anal. Chem. 80 (2008) 461–473. doi:10.1021/ac071504j.
[16] W.C. Siegler, B.D. Fitz, J.C. Hoggard, R.E. Synovec, Experimental study of the quantitative precision for valve-based comprehensive two-dimensional gas chromatography, Anal. Chem. 83 (2011) 5190–5196. doi:10.1021/ac200302b.
[17] M.S. Klee, J. Cochran, M. Merrick, L.M. Blumberg, Evaluation of conditions of comprehensive two-dimensional gas chromatography that yield a near-theoretical maximum in peak capacity gain, J. Chromatogr. A 1383 (2015) 151–159. doi:10.1016/j.chroma.2015.01.031.
[18] N.E. Watson, W.C. Siegler, J.C. Hoggard, R.E. Synovec, Comprehensive three-dimensional gas chromatography with parallel factor analysis, Anal. Chem. 79 (2007) 8270–8280. doi:10.1021/ac070829x.
[19] N.E. Watson, H.D. Bahaghighat, K. Cui, R.E. Synovec, Comprehensive three-dimensional gas chromatography with time-of-flight mass spectrometry, Anal. Chem. 89 (2017) 1793–1800. doi:10.1021/acs.analchem.6b04112.
[20] N.E. Watson, S.E. Prebihalo, R.E. Synovec, Targeted analyte deconvolution and identification by four-way parallel factor analysis using three-dimensional gas chromatography with mass spectrometry data, Anal. Chim. Acta. 983 (2017) 67–75. doi:10.1016/j.aca.2017.06.017.
[21] W.C. Siegler, J.A. Crank, D.W. Armstrong, R.E. Synovec, Increasing selectivity in comprehensive three-dimensional gas chromatography via an ionic liquid stationary phase column in one dimension, J. Chromatogr. A 1217 (2010) 3144–3149. doi:10.1016/j.chroma.2010.02.082.
[22] C.G. Fraga, Chemometric approach for the resolution and quantification of unresolved peaks in gas chromatography–selected-ion mass spectrometry data, J. Chromatogr. A 1019 (2003) 31–42. doi:10.1016/S0021-9673(03)01329-3.
[23] S.E. Stein, An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data, J. Am. Soc. Mass Spectrom. 10 (1999) 770–781. doi:10.1016/S1044-0305(99)00047-1.
[24] C.G. Fraga, C.A. Bruckner, R.E. Synovec, Increasing the number of analyzable peaks in comprehensive two-dimensional separations through chemometrics, Anal. Chem. 73 (2001) 675–683. doi:10.1021/ac0010025.
[25] J.B. Callis, Super resolution in chromatography by numerical deconvolution, in: ultrahigh resolution chromatography, ACS, 1984: pp. 171–198. doi:10.1021/bk-1984-0250.ch012.

[26] A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, Trilinear chemometric analysis of two-dimensional comprehensive gas chromatography–time-of-flight mass spectrometry data, J. Chromatogr. A 1027 (2004) 269–277. doi:10.1016/j.chroma.2003.08.081.
[27] S.E. Prebihalo, K.L. Berrier, C.E. Freye, H.D. Bahaghighat, N.R. Moore, D.K. Pinkerton, R.E. Synovec, Multidimensional gas chromatography: advances in instrumentation, chemometrics, and applications, Anal. Chem. 90 (2018) 505–532. doi:10.1021/acs.analchem.7b04226.
[28] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta. 185 (1986) 1–17. doi:10.1016/0003-2670(86)80028-9.
[29] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chemometr. Intell. Lab. Syst. 2 (1987) 37–52. doi:10.1016/0169-7439(87)80084-9.
[30] Richard Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Wiley, 2003.
[31] K. Beebe, Chemometrics: A Practical Guide, Wiley, 1998.
[32] N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: An overview, Talanta 123 (2014) 186–199. doi:10.1016/j.talanta.2014.02.003.
[33] R.A. Fisher, Statistical Methods for Research Workers, 14th ed., Oliver and Boyd, 1970.
[34] T. Gröger, M. Schäffer, M. Pütz, B. Ahrens, K. Drew, M. Eschner, R. Zimmermann, Application of two-dimensional gas chromatography combined with pixel-based chemometric processing for the chemical profiling of illicit drug samples, J. Chromatogr. A 1200 (2008) 8–16. doi:10.1016/j.chroma.2008.05.028.
[35] K.M. Pierce, J.C. Hoggard, J.L. Hope, P.M. Rainey, A.N. Hoofnagle, R.M. Jack, B.W. Wright, R.E. Synovec, Fisher ratio method applied to third-order separation data to identify significant chemical components of metabolite extracts, Anal. Chem. 78 (2006) 5068–5075. doi:10.1021/ac0602625.
[36] L.M. Dubois, K.A. Perrault, P.-H. Stefanuto, S. Koschinski, M. Edwards, L. McGregor, J.-F. Focant, Thermal desorption comprehensive two-dimensional gas chromatography coupled to variable-energy electron ionization time-of-flight mass spectrometry for monitoring subtle changes in volatile organic compound profiles of human blood, J. Chromatogr. A 1501 (2017) 117–127. doi:10.1016/j.chroma.2017.04.026.
[37] L.C. Marney, W. Christopher Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher-ratio software for improved feature selection analysis of comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry data, Talanta 115 (2013) 887–895. doi:10.1016/j.talanta.2013.06.038.
[38] B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher ratio analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC–TOFMS) data using a null distribution approach, Anal. Chem. 87 (2015) 3812–3819. doi:10.1021/ac504472s.
[39] B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, Chemical characterization of the acid alteration of diesel fuel: Non-targeted analysis by two-dimensional gas chromatography coupled with time-of-flight mass spectrometry with tile-based Fisher ratio and combinatorial threshold determination, J. Chromatogr. A 1440 (2016) 179–190. doi:10.1016/j.chroma.2016.02.067.
[40] N.E. Watson, B.A. Parsons, R.E. Synovec, Performance evaluation of tile-based Fisher Ratio analysis using a benchmark yeast metabolome dataset, J. Chromatogr. A 1459 (2016) 101–111. doi:10.1016/j.chroma.2016.06.067.

[41] A. Eftekhari, H. Parastar, Multivariate analytical figures of merit as a metric for evaluation of quantitative measurements using comprehensive two-dimensional gas chromatography–mass spectrometry, J. Chromatogr. A 1466 (2016) 155–165. doi:10.1016/j.chroma.2016.09.016.
[42] R. Bro, PARAFAC. Tutorial and applications, Chemometr. Intell. Lab. Syst. 38 (1997) 149–171. doi:10.1016/S0169-7439(97)00032-4.
[43] D.K. Pinkerton, B.A. Parsons, T.J. Anderson, R.E. Synovec, Trilinearity deviation ratio: A new metric for chemometric analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Anal. Chim. Acta. 871 (2015) 66–76. doi:10.1016/j.aca.2015.02.040.
[44] M. Navarro-Reig, J. Jaumot, T.A. van Beek, G. Vivó-Truyols, R. Tauler, Chemometric analysis of comprehensive LC×LC-MS data: Resolution of triacylglycerol structural isomers in corn oil, Talanta 160 (2016) 624–635. doi:10.1016/j.talanta.2016.08.005.
[45] M.L. Oca, M.C. Ortiz, A. Herrero, L.A. Sarabia, Optimization of a GC/MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware, Talanta 106 (2013) 266–280. doi:10.1016/j.talanta.2012.10.086.
[46] J.C. Hoggard, R.E. Synovec, Automated Resolution of Nontarget Analyte Signals in GC × GC-TOFMS Data Using Parallel Factor Analysis, Anal. Chem. 80 (2008) 6677–6688. doi:10.1021/ac800624e.

Chapter 2. Comprehensive Two-Dimensional Gas Chromatography Using Partial Modulation via a Pulsed Flow Valve with a Short Modulation Period

This chapter was reproduced from C.E. Freye, H.D. Bahaghighat, R.E. Synovec, "Comprehensive Two-Dimensional Gas Chromatography Using Partial Modulation via a Pulsed Flow Valve with a Short Modulation Period" *Talanta* 177 (2018) 142–149.

2.1 INTRODUCTION

Comprehensive two-dimensional (2D) gas chromatography (GC × GC) is a powerful technique for the analysis of complex samples containing analytes that are naturally volatile, semi-volatile, or amenable to gas phase analysis after derivatization [1–5]. Addition of a secondary separation dimension, ^{2}D, provides a 5-fold to 15-fold increase in total peak capacity over a one-dimensional (1D) separation (1D-GC). In order to transfer analytes from the primary ^{1}D column to the ^{2}D column, a modulation interface must be present. The modulator operates by injecting portions of the ^{1}D column effluent onto the ^{2}D column with a user defined, relatively rapid frequency. There are many types of modulators in GC × GC, but all fall under the three basic categories: thermal [1,6–15], valve-based [16–21], and flow modulation [22–29].

Thermal modulation relies on a low temperature to trap and focus analytes as they elute from the ^{1}D column and introduces them in a narrow pulse to the ^{2}D column through rapid heating. There are three types of thermal modulators: resistively heated trap [1,6,7], heated sweeper [8,9], and cryogenic focus, which is often divided into longitude movable trap [10] and jet trap [11–15]. Phillips and Liu developed the first thermal modulator using a slotted heater which swept across a thick-filmed capillary connecting the ^{1}D column to the ^{2}D column [1]. Analytes eluting from

the ^{1}D column are retained in the thick film and then volatilized and injected onto the ^{2}D column when the heater sweeps over the capillary. The most common thermal modulator uses a cryogen to trap analytes before reinjecting them. Marriott and Kinghorn [10] pioneered the longitude movable cryogen trap, whereby a jet of fluid CO_2 is applied to produce a moving cryogenic trap with a small portion of capillary between the two columns. The condensed analytes are injected onto the ^{2}D column when the capillary is rapidly heated. Another cryogenic modulator, and perhaps the most commonly used thermal modulator, is the jet trap developed by Ledford and Billesbach [11], which uses strategically placed and timed jets of cryogenic gas or a combination of heat and cooled jets. Numerous improvements have been made to these thermal modulators.

Valve-based modulation commonly employs a diaphragm valve to divert effluent from the ^{1}D to ^{2}D column [16]. The diaphragm valve utilized in the initial report had a temperature limit of 175 °C due to the temperature sensitive O-rings, and only ~ 10% of the effluent from ^{1}D column was injected onto the ^{2}D column. Improvements by Seeley increased detection sensitivity by employing a sample loop to collect the effluent from the ^{1}D separation before injection onto the ^{2}D column, with the fraction of the ^{1}D effluent transferred to the ^{2}D column approaching 100% [17]. Using a significantly higher flow rate on the ^{2}D column compared to the ^{1}D column, sample pulses are compressed within the sample loop creating narrow ^{2}D peaks while improving detection limits compared to 1D-GC [17,18]. The original temperature limit of 175 °C was overcome by face mounting the valve on the top of the GC, which extended the temperature limit to 265 °C [19]. Recent advances in diaphragm valve technology have extended the temperature limit to 325 °C with the valve mounted directly in the oven, avoiding the complication and shortcomings of face mounting the valve [18]. Diaphragm valve-based modulation has also been demonstrated to be compatible with mass spectrometry [20,21].

Flow modulation has its origins in Deans switching which is often used in heart-cutting GC-GC to transfer a portion of the effluent from the first GC column onto a subsequent GC column [22]. A Deans switch has been adapted by Seeley for use as a modulator for GC × GC [23]. Differential flow modulation is another type of flow modulation which closely resembles Deans switching, but results in total transfer of effluent from the ^{1}D column to the ^{2}D column. Differential flow modulation involves the alternation of filling and flushing cycles of a collection channel controlled by a solenoid valve [24]. Flow modulation is gaining popularity due to the operational simplicity, reliability, and capability to modulate components over a wide volatility range [22–28]. Flow modulation has also been demonstrated in an atypical manner. A pattern of the primary effluent is transferred to the secondary column instead of single pulses [29]. The detected signal is then deconvoluted and transformed into a GC × GC chromatogram. While extra processing is required to convert the detected signal into more readily interpretable data, more intense and narrower peaks are obtained compared to a traditional flow modulation technique.

Another interesting, yet unconventional, modulation approach was introduced by Cai and Stearns [30], a method referred to as "partial modulation." Using a custom built pulse flow modulator, small pulses of carrier gas are repetitively injected at the interface between the ^{1}D and ^{2}D columns, creating either local high or low concentration pulses in the effluent departing the ^{1}D column, which are then separated on the ^{2}D column. Dependent upon the mode of modulation, either positive or negative "spikes" in analyte concentration are created. When the injected pulses are negative, the resulting ^{2}D separations are essentially performing vacancy chromatography. A modulation period, P_M, of 1 s was applied in this previous report. Since partial modulation produced raw data in an unconventional format, additional data processing is required to achieve a conventional appearing 2D plot of the GC × GC separation. The ^{2}D peak widths

observed were ${}^2W_b \sim 50$ ms and larger [30], similar to what is achieved by thermal [10–15], diaphragm valve [16–21,31], and standard flow modulation [23–28].

Partial modulation lends itself to producing a short P_M because the effluent from the ^{1}D column does not have to be cryogenically trapped or collected in a sample loop or channel. Moreover, there is no desorption or injection period, which facilitates very narrow peaks on the ^{2}D separation and thus a high ^{2}D peak capacity. While the initial report using partial modulation method was intriguing [30], there is untapped potential for producing extremely narrow ^{2}D peaks with ${}^2W_b$ smaller than 50 ms, and extremely short P_M. Being able to do so may open up new opportunities in the development of novel high speed versions of multi-dimensional GC-based separation technology involving "GC-sensors" [31–35].

In order to begin to explore these high speed GC-based technologies, herein we study the feasibility of a pulse flow valve modulator for GC × GC, implementing partial modulation using a commercially available pulse flow valve. Very short modulation periods are examined, P_M of 50, 100, 250 and 500 ms, with 50 and 100 ms notably shorter than the smallest P_M reported of 200 ms [31]. Using the commercially available pulse flow valve, this injection technique will be shown to perform a combination of vacancy chromatography and frontal analysis, whereby each pulse disturbance in the analyte concentration departing the ^{1}D column results in data that is readily converted into a ^{2}D separation. A procedure is introduced to convert the raw data obtained into the 2D form typical of a GC × GC separation.

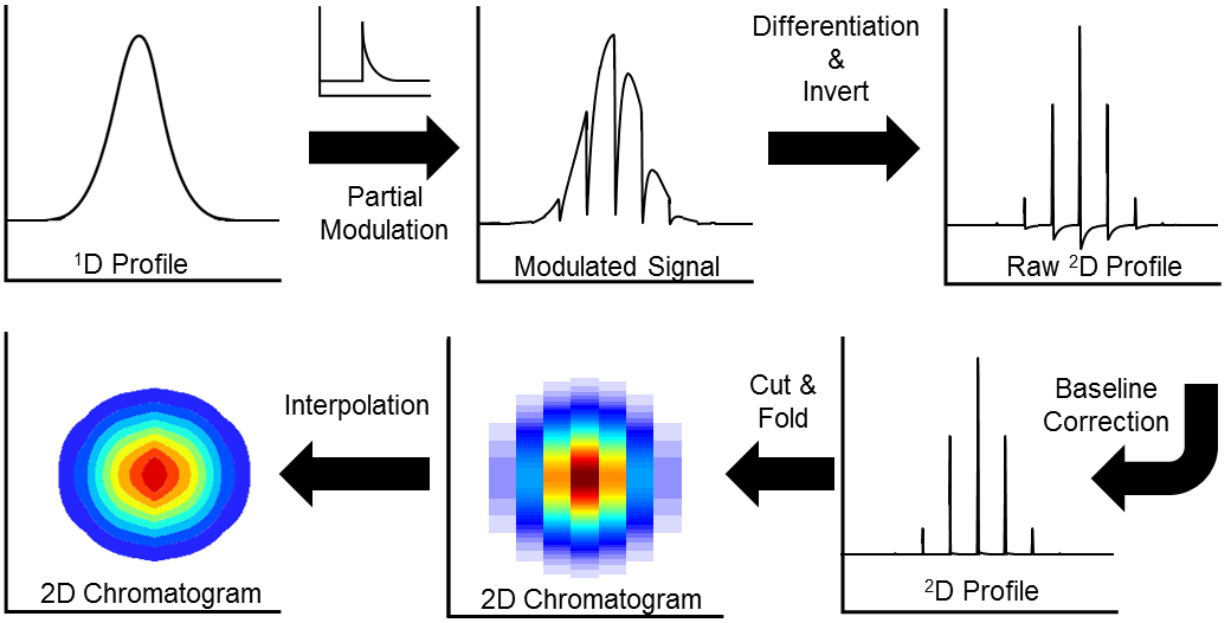


Figure 2.1 Modulation process using the pulse flow valve and conversion of the modulated signal into two-dimensional (2D) format with the appearance of a GC × GC chromatogram.

The basic concept is illustrated in Figure 2.1, which shows the modulation of a single analyte using the commercial pulse flow valve and the subsequent conversion of the raw data into the typical 2D form of an analyte signal in a GC × GC chromatogram. The pulse flow valve creates pressure pulses in the shape of an exponentially modified Gaussian where the leading edge of each pressure pulse is extremely sharp and then slowly decays to zero. Negatively modulated signals are created as in vacancy chromatography, while also containing basic elements of a frontal analysis measurement. The raw data is differentiated to convert it into a "traditional" 1D raw vector plot. Differentiation of the sharp leading edge of the analyte signal creates the resulting ^{2}D peak, while differentiation of the slowly decaying portion of the analyte signal results in a modest baseline sag that is readily removed using a baseline correction algorithm. The baseline corrected data is then

folded at the modulations to produce the typical 2D form of the GC × GC chromatogram. The experimental data and resulting GC × GC chromatograms will be shown to support the illustration in Figure 2.1. For this study, a separation of a 115-component test mixture with a wide range of boiling points (36–372 °C) of nine different sample classes is evaluated using P_M ranging from 50 to 500 ms.

2.2 EXPERIMENTAL

The GC × GC instrument with the pulse flow valve modulator was evaluated using a flame ionization detector (FID). The instrumental platform consisted of an Agilent 6890 GC (Agilent Technologies, Palo Alto, CA, USA) modified in-house with high-speed pulse valve model 009–1643-900 (Parker Hannifin, Hollis, NH, USA) mounted outside the oven. The pulse flow valve was controlled using LabVIEW 8.2 (National Instruments, Austin, TX, USA) interfaced to a National Instruments data acquisition board. The stock electrometer for the Agilent FID was

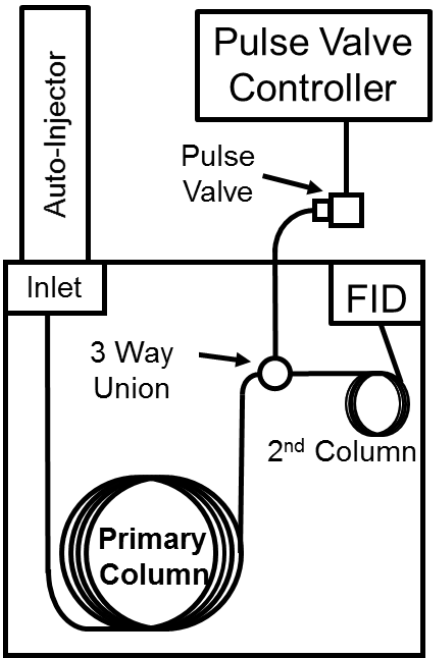


Figure 2.2 Instrumental schematic to perform comprehensive 2D gas chromatography (GC × GC) using partial modulation via the pulse flow valve.

replaced with a high-speed electrometer built in-house allowing the data to be collected at 100 kHz, with the data boxcar averaged to 1 kHz. The electrometer was interfaced to a National Instruments data acquisition board, and the resulting data was collected using an in-house written LabVIEW program. Post-run data processing was performed in Matlab R2015b (The Mathworks, Inc., Natick, MA, USA). A schematic of the instrument is shown in Figure 2.2.

Samples were introduced to the GC × GC via a 7683B auto-injector (Agilent Technologies, Palo Alto, CA, USA). The inlet and FID temperature were both set to 250 °C. The ^{1}D column was a RTX-5ms (diphenyl dimethyl polysiloxane) stationary phase: 40 m length, 180 µm inner diameter (i.d.), and 0.4 µm film thickness. The ^{2}D column was a RTX-Wax (polyethylene glycol) stationary phase: 2.0 m length, 180 µm i.d., and 0.180 µm film thickness. ^{1}D and ^{2}D columns were linked using a 3-way T-union model MT.5CXS6 (Valco Instruments Company Inc., Houston, TX, USA). An in-house fitting was fabricated to mate the pulse flow valve to 7.24 cm × 1.65 mm copper tubing (Restek, Bellefonte, PA, USA) reduced to 3.81 cm × 0.635 mm steel tubing interfaced with a 2 µl sample loop model CSL2 (Valco Instruments Company Inc.) which was connected to the 3-way T-union. Prior to each sample injection, HPLC grade hexane and acetone obtained from Fisher Scientific were used as solvent rinses. Ultra-high purity hydrogen (Grade 5, 99.999%) was used as the carrier gas (Praxair, Seattle, WA, USA) at a constant flow of 4 ml/min. Oven temperature control on ^{2}D was not applied, so both columns were at the same temperature.

A test mixture containing a diverse set of 115 compounds was used to evaluate the instrumental platform. The test mixture contained a wide range of boiling points (36–372 °C) with nine chemical compound classes: alkanes, esters, halogenated alkanes, aromatics, alkenes, alkynes, ketones, alcohols, cycloalkanes [35]. A 1 µl volume of the test mixture was injected at a split of 240:1. For each test mixture separation, the oven was held at 40 °C for 1.5 min and ramped

at 40 °C/min to 220 °C where it was held for 3 min. Various modulation periods were studied: P_M of 50, 100, 250, and 500 ms, all using an injection pulse width of 2 ms (i.e. how long the pulse valve injects carrier gas). Using the Bernoulli equation, it is estimated that ~ 11 µl of carrier gas is injected for each 2 ms modulation. The head pressure on the valve was held at 32 psig for 1.5 min and ramped at 4.0 psig/min to 50 psig and held for 3 min. The head pressure on the pulse valve must be higher than the exit pressure of the ^{1}D column, otherwise the analytes will not be interrupted by the carrier gas injection from the pulse flow valve. The above mentioned pressure was determined to provide suitable flow to interrupt the shape of the ^{1}D profiles enabling the data processing to be performed. However, care must be taken to not overpressure the pulse valve as it will cause increased retention times on the ^{1}D dimension, ^1t_R.

The suitability of pulse flow valve modulation for quantitative analysis was investigated. Two solutions, each 5.2 ppt (parts per thousand by weight) of cyclohexane and benzene diluted in dicholoromethane, were quantitatively combined to form seven calibration solution mixtures of benzene to cyclohexane by percent weight of each of these two solutions: 100–0%, 90–10%, 75–25%, 50–50%, 25–75%, 10–90%, and 0–100%. A 0.1 µl volume of each solution mixture was injected at a split of 80:1. Three replicates of each solution mixture were obtained using a ^{1}D RTX-5ms (diphenyl dimethyl polysiloxane) 8 m length × 100 µm i.d. × 0.1 µm film thickness and a ^{2}D RTX-Wax 1.0 m length, 100 µm i.d., and 0.1 µm film thickness. A modulation period of 250 ms was used and the oven was set to 65 °C with a column flow of 3.0 ml/min and a head pressure on the pulse flow valve of 80 psig. All other instrumental parameters were the same as mentioned above.

2.3 RESULTS AND DISCUSSION

The GC × GC – FID chromatogram for the test mixture using a modulation period P_M of 250 ms in the raw 1D vector form is shown in Figure. 2.3A. A representative analyte is shown in Figure 2.3B (box indicated in Fig. 2.3A) which is relatively unretained on ^{2}D has been selected to demonstrate the data processing (to convert the raw data into a format with the appearance of a GC × GC chromatogram). The inset in Fig. 2.3B shows a single "peak" created by the modulator. The sharp decrease in signal represents the ^{2}D analyte response of interest, nominally in the form of an "error" function, followed by a slow decay of the analyte response back to the original ^{1}D concentration profile. This ^{2}D analyte response is created by rapid actuation of the pulse flow valve in combination with the three-way union and coupled with the tubing assembly (see Fig. 2.2). Reducing the tubing assembly from a large inner diameter to a smaller inner diameter results in a pressure wave in the form of an exponentially modified Gaussian with a very sharp leading edge. This pressure wave causes the analytes to have the "error" function profile followed by a slow decay back to the original ^{1}D concentration profile. After the raw vector data is baseline corrected and smoothed, the differential of the raw vector data is taken and is then multiplied by negative one to invert the chromatogram creating peaks that have positive intensities as shown in Fig. 2.3C. The differentiation process, defined as ΔSignal/ΔTime, was calculated by taking the difference between two adjacent signal intensity values (FID Signal) and dividing by the time interval (1 ms). By taking the differential of the data, each "error" function signal response is nominally turned into a Gaussian-like analyte peak. The width of each ^{2}D Gaussian peak is dependent on the slope of the error function, and ^{2}D the retention time is defined by the inflection point of the error function. However, due to the decay of the analyte concentration back to the ^{1}D concentration profile, a sag in baseline is initially observed following the differentiation step. The baseline sag is readily removed using a rolling minimum baseline correction program written in-house [36],

which objectively removes the low frequency drift in the baseline, with the result shown in Fig. 2.3D. The baseline correction step was found to not adversely affect the shape nor areas of the ^{2}D peaks, as is demonstrated herein. While the differentiation and baseline correction steps must be taken to visualize the data in the form of a traditional GC × GC chromatogram, useful chemical information is inherently encoded within the data when it is in the raw non-differentiated form.

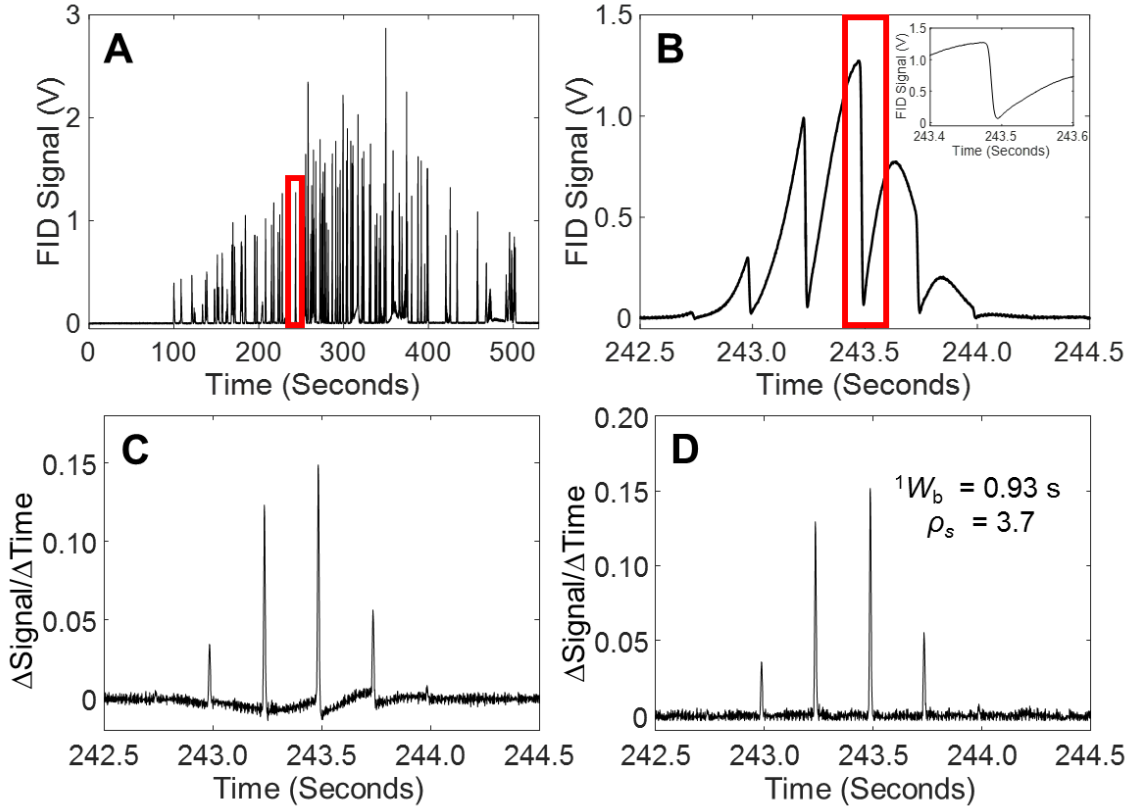


Figure 2.3 (A) Chromatogram (raw data on ^{1}D) of the 115 component test mixture collected using a modulation period P_M of 250 ms. (B) Shown is a mid-elution non-overlapped analyte on ^{1}D. The inset figure shows a zoom of a single "peak" create by pulse modulation. (C) Processed data wherein the raw data in (B) has been differentiated to produce ^{2}D peaks. (D) Data from (C) has been baseline corrected using a rolling minimum. Peaks on the ^{2}D separation are narrow yielding a high peak capacity despite a short modulation period of P_M of 250 ms.

Additional figures-of-merit for the representative analyte are discussed using the processed 1D vector form as shown in Fig. 2.3D. The apparent width on the first dimension, ${}^{1}W_{b}$ is 0.93 s, determined by fitting a Gaussian function through the apexes of each of the ^{2}D peaks in the 1D vector format [18,21,37–39]. With a modulation period P_M of 250 ms, the apparent sampling density across the ^{1}D peak width-at-base ($4{}^{1}\sigma$), defined as $\rho_S = {}^{1}W_b / P_M$, was experimentally

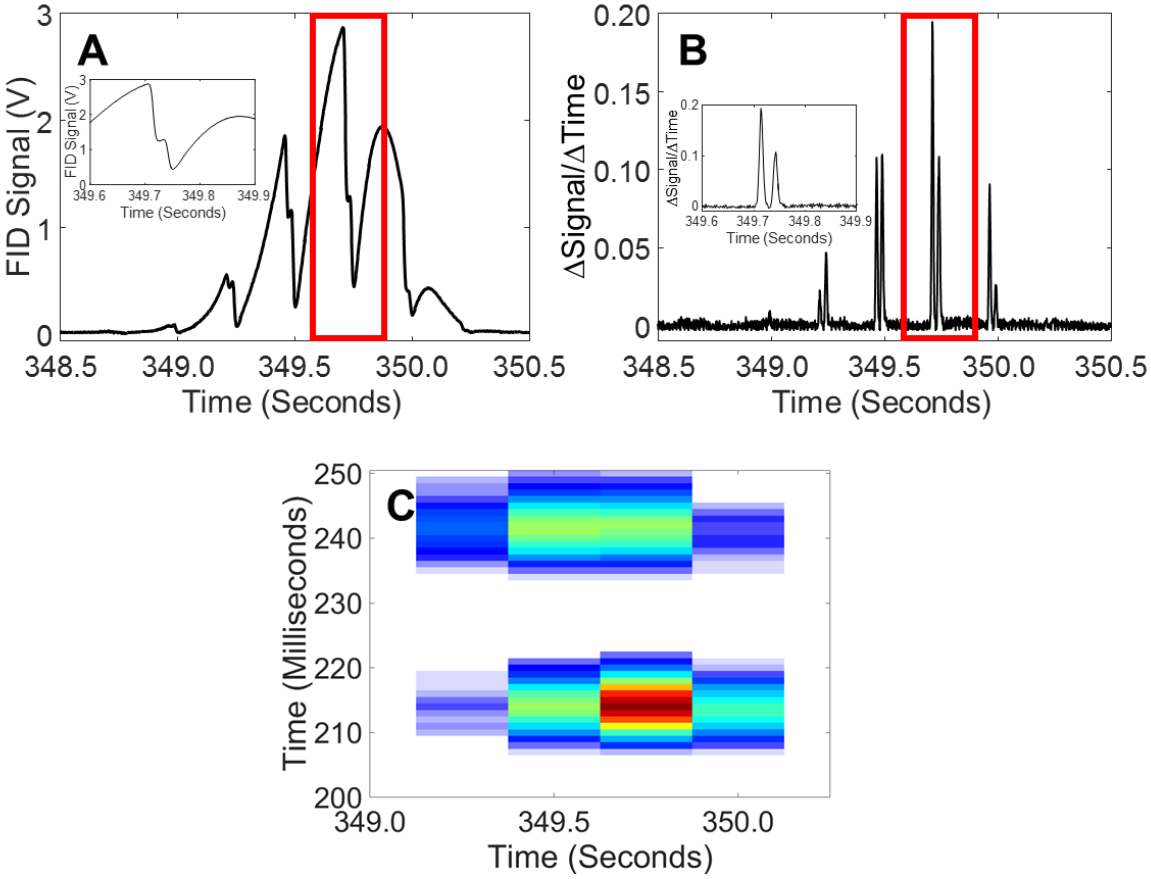


Figure 2.4 An example of an overlapped peak on the ^{1}D dimension from Fig. 2.3A for a modulation period P_M of 250 ms. (A) Raw data on ^{1}D for two overlapped analytes. The inset shows an enhanced view, for one modulation. (B) After data processing, the two analytes are transformed into Gaussian-like peaks which are resolved on the ^{2}D dimension. (C) View of the fully processed data in a GC × GC separation format.

determined to be 3.7 [40]. The sampling density, ρ_S, defined as such is equivalent to the modulation ratio, M_R [38], but ρ_S is more descriptive of the relationship between the data sampling rate (modulation frequency) and the ^{1}D peak width so is used throughout. For either valve, flow or thermal modulation, a $\rho_S < 2.0$ is deemed to only be semi-quantitative and can lead to high %RSD for measured peak areas [23,38,39]. The average apparent ^2W_b was an impressive 13 ms for this representative analyte in Fig. 2.3D, which is significantly more narrow than any previous reported ^2W_b for conventional GC × GC separations.

We now demonstrate the data processing steps, as illustrated in Fig. 2.1, for a pair of analytes that co-elute in the ^{1}D separation from Fig. 2.3, but are separated in the ^{2}D separation. Fig. 2.4A shows the raw 1D vector plot of the data of two analytes that are overlapped in the ^{1}D separation. The inset in Fig. 2.4A shows a single modulation containing the responses for the two analytes. In the sharp decreasing portion of the single modulation, there are two inflection points, one for each analyte, located at a different times indicating the analytes are separated on the ^{2}D column. Following the data processing steps as presented in Fig. 2.3B-D, the resulting two analyte peaks are evident in the context of the ^{1}D and ^{2}D dimensions in Fig. 2.4B. The two analytes have different apparent ^2W_b peak widths in the resulting ^{2}D peaks (Fig. 2.4B inset), a direct result of having different slopes for their error functions (Fig. 2.4A inset). As expected, the analyte which is retained less (lower ^2k') has a smaller ^2W_b. Fig. 2.4C shows the 2D view, in the form of a GC × GC chromatogram, of the fully processed data for the analytes overlapped on ^{1}D.

Following the data processing steps in Fig. 2.3B-D, the raw data from Fig. 2.3A has been converted into the GC × GC – FID chromatogram of the test mixture using a P_M of 250 ms, as presented in Fig. 2.5A. The chromatogram has been re-registered by 680 ms on ^{2}D to remove the dead time offset. With the re-registered format, the earliest analyte elutes at 34 ms on ^{2}D, while the most retained analyte elutes at 265 ms (wrapped around into the next modulation). Based upon the re-registration time of 680 ms, the retention factors on ^{2}D range from $^2k'$ ~ 0.05–0.40. These small $^2k'$ values benefit the higher efficiency (narrow chromatographic peaks) obtained on the ^{2}D separation [31,32,41]. Due to the mismatch of the pressure gradient from the pulsed flow valve and the exit of ^{1}D exit pressure, the dead time on ^{2}D changes ~ 40 ms across the entire ^{1}D separation. This is visually manifested by the change in the ^2t_R of the alkanes (lowest $^2k'$). Common implementation of this technique will require a better understanding of how to address

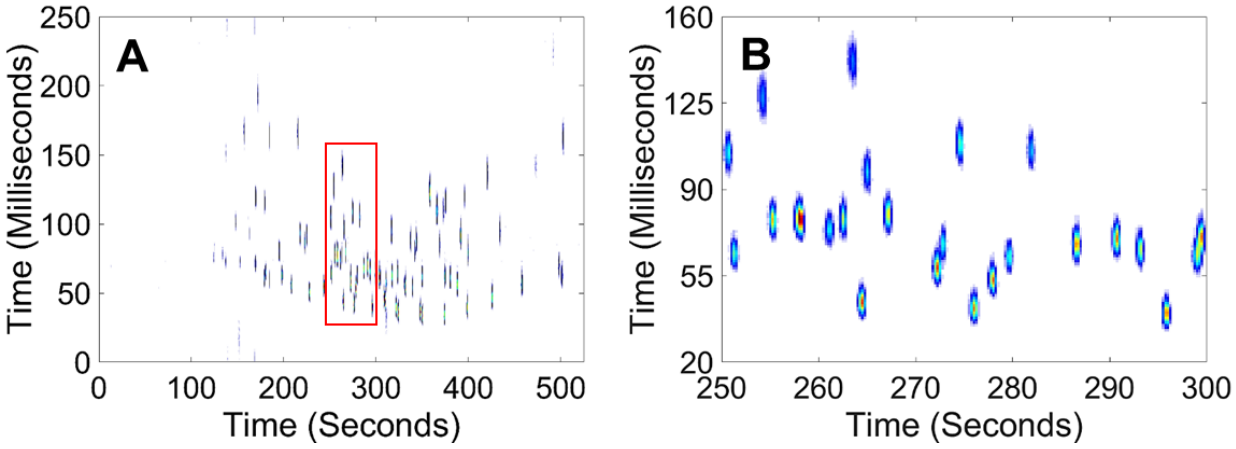


Figure 2.5 (A) GC × GC – FID chromatogram of the 115 component test mixture collected using a modulation period P_M of 250 ms. Fully processed data from Fig. 2.3A in the form of Fig. 2.4C. The ^{2}D separation has been reregistered by 680 ms to remove the dead time. (B) Enhanced view of a selected portion of the chromatogram is shown.

these pressure and flow matching challenges. To better assess the performance of this separation approach, an enhanced portion (box indicated in Fig. 2.5A) of a portion of the separation is shown in Fig. 2.5B. Remarkably, the ^{2}D apparent peak widths in this time window range from 15 ms to 25 ms.

In order to discuss the GC × GC separation peak capacity achieved, the relationship between apparent ^2W_b and retention time ^2t_R in the ^{2}D separation dimension was studied. For this purpose, data from separations using a P_M of 50, 100, 250, and 500 ms were collected, although only the P_M of 100, 250, and 500 ms are compared in Fig. 2.6 since no analytes were wrapped around into lesser retained analytes from the subsequent modulation. As shown in Fig. 2.6, pulse flow valve modulation provides a linear relationship between apparent ^2W_b and retention time ^2t_R, as expected [42–44]. However, it is interesting to note for a given analyte, its ^2t_R directly correlates to the P_M applied. That is, the smaller P_M produced a smaller ^2t_R for the same analyte, with the effect more pronounced for the more retained analytes. For example, diethyl phthalate, the most retained analyte, has a ^2t_R of 323 ms for a P_M of 500 ms, ^2t_R of 262 ms for a P_M of 250 ms,

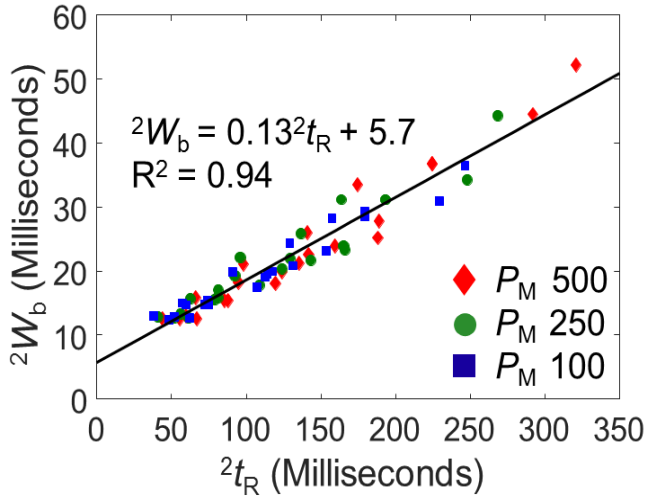


Figure 2.6 Plot of apparent ^2W_b versus ^2t_R for a set of selected analytes for three different modulation periods P_M of 500, 250 and 100 ms. ^2t_R has been adjusted to remove the dead time.

and ${}^2t_R$ of 245 ms for a P_M of 100 ms. 2-decanone, a moderately retained analyte, has a ${}^2t_R$ of 158 ms for a P_M of 500 ms, a ${}^2t_R$ of 142 ms for a P_M of 250 ms, and a ${}^2t_R$ of 137 ms for a P_M of 100 ms. Hexadecane, the least retained analyte, has ${}^2t_R$ of 39 ms, 36 ms, and 33 ms for P_M of 500 ms, 250 ms, and 100 ms, respectively. Additionally, as the modulation period P_M was decreased, analytes were more retained on ^{1}D (Δ^1t_R of ~ 30 s from P_M of 500–50 ms). Both these observations are attributed to the pulse flow valve actuation injecting carrier gas at an increased rate which delays elution on the ^{1}D separation but drives analytes through the ^{2}D column at a faster rate. Additionally, ${}^1W_b$ also increases but at an insignificant amount (Δ^1W_b of ~ 90 ms from a P_M of 500–50 ms). This phenomenon could be due to the delayed ${}^1t_R$, and/or due to the analyte being broadened by the modulator during the transfer from the ^{1}D column to the ^{2}D column. The P_M of 50 ms was the fastest modulation period observed whereby the pulse flow valve was found to reliably and reproducibly modulate.

Based on the fitted equation in Fig. 2.6, for a P_M of 250 ms allowing no wrap around and full use of the ^{2}D space, the peak capacity on ^{2}D was determined to be ${}^2n_c$ = 10. Increasing the P_M to 500 ms only increased ${}^2n_c$ by a modest amount to 12 due to the nonlinear relationship between ${}^2n_c$ and ${}^2k$' (which is linearly related to apparent ${}^2W_b$) [42]. For the ${}^2n_c$ determination, a simple peak capacity equation, i.e. ${}^2n_c = P_M/{}^2W_{b,avg}$, where ${}^2W_{b,avg}$ is the average apparent peak width on ^{2}D, is not sufficient due to the pseudo-isothermal ^{2}D separations. Using the ${}^2W_b$ of an unretained analyte to calculate ${}^2n_c$ can also result in overinflated values. Therefore, the following recursive equation was applied to accurately determine ${}^2n_c$,

$$R_S = \frac{t_{R,n} - t_{R,n-1}}{0.5(W_{b,n} - W_{b,n-1})} = 1 \qquad (2.17)$$

The relationship between retention time and ${}^2W_b$ was discussed in the previous paragraph (and illustrated in Fig. 2.6). Substituting the equation from Fig. 2.6 into Eq. (2.1) for both $W_{b,n}$ and

$W_{b,n-1}$ and rearranging allows the calculation of the retention times of successive peaks (and R_s = 1) based on the measured dead time, $t_{R,(n=0)}$ on ^{2}D. These retention times were used to calculate apparent ^2W_b from the experimentally observed linear relationship between ^2t_R and apparent ^2W_b. Individual Gaussian peaks of equivalent area were modeled based on these peak widths and retention times, then concatenated to give a continuous chromatogram [42,43]. The number of peaks that fit in a given P_M provides the accurate determination of ^2n_c. Finally, the ideal total peak capacity, $n_{c,2D}$, defined as $^1n_c \times {}^2n_c$, was calculated assuming ^1W_b of 0.93 s for a run time of 400 s

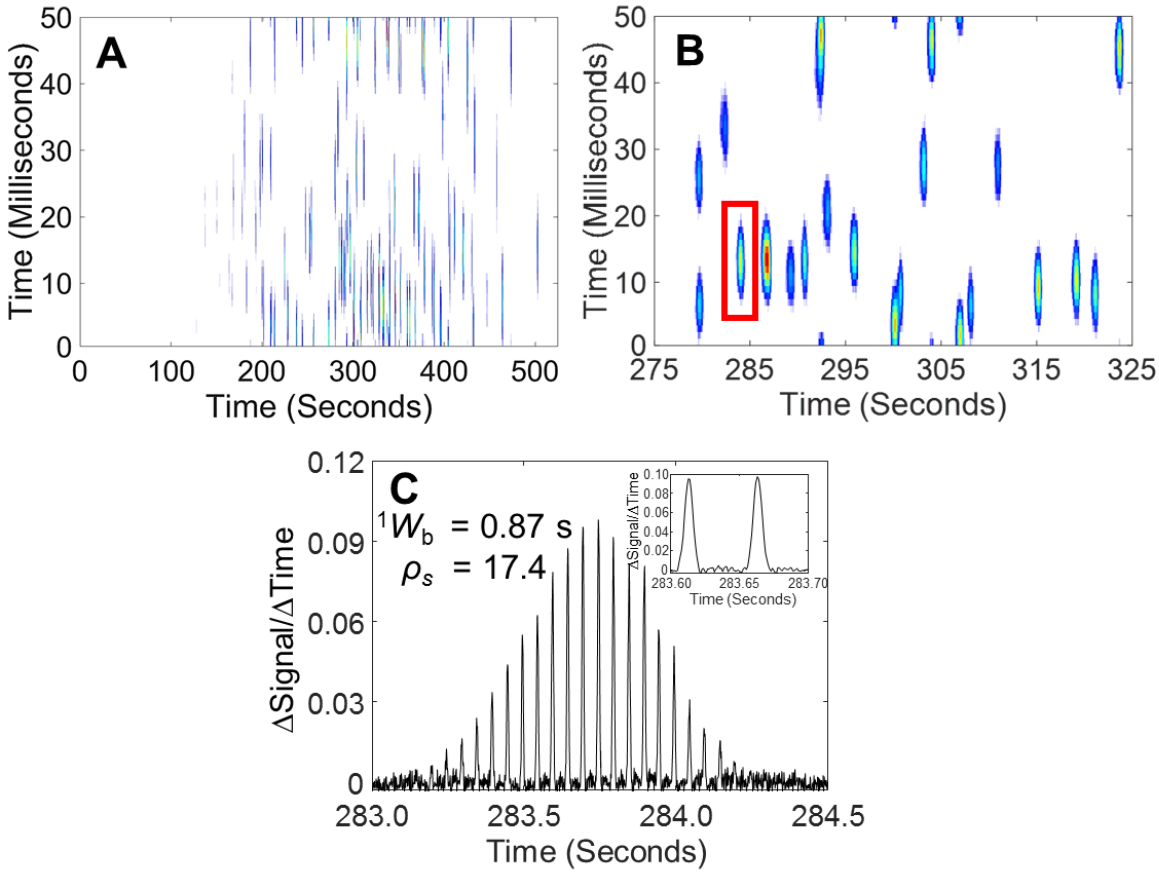


Figure 2.7 (A) GC × GC – FID chromatogram, as in Fig. 2.5, using a modulation period P_M of 50 ms instead of 250 ms. (B) Enhanced view of a selected portion of the chromatogram is shown. (C) Raw 1D vector form of the GC × GC – FID data for a single peak, highlighted in (B). Due to the ultra-fast P_M, the peak is highly sampled. The insert shows the peaks for two modulations.

(first analyte elutes at ${}^1t_R$ = 100 s and last analyte at ${}^1t_R$ = 500 s). Therefore, the $n_{c,2D}$ (${}^1n_c$ ~ 430) × (${}^2n_c$ ~ 10) is ~ 4300 or a peak capacity production of ~ 650 peaks/min for a P_M of 250 ms.

Using the modulation period P_M of 50 ms highlights the ability of pulse flow valve modulation to produce extremely fast GC × GC separations, demonstrating the enabling potential of this technology for high speed, high peak capacity GC-sensors [32–35]. Fig. 2.7A shows the GC × GC – FID chromatogram of the test mixture using a P_M of 50 ms after the data processing steps (Fig. 2.3B-D). The results for a P_M of 500 and 100 ms, while summarized in Fig. 2.6 and elsewhere, are not shown in the same detail as Fig. 2.7 for brevity. The chromatogram in Fig. 2.7A has been re-registered on ^{2}D to remove the dead time offset. Fig. 2.7B shows an enhanced portion of Fig. 2.7A in order to discuss merits of the short P_M of 50 ms. Due to the short P_M, analytes are wrapped around into the next modulation, however based upon the narrow apparent ${}^2W_b$ of ~ 12 ms, a ${}^2n_c$ ~ 4 is readily estimated. Fig. 2.7C shows the raw vector form of a moderately retained analyte. The insert shows the two most intense modulated ^{2}D peaks. Due to the ultra-fast P_M, the analyte ^{1}D peak is highly sampled with a ρ_S of 17.4. Future application of the 50 ms P_M would involve a ρ_S of 2–4, adequate for sampling ^{1}D peaks in GC × GC separations. Hence, it is intriguing to consider application of the pulse valve in conjunction with a modified injection interface to provide narrower ^{1}D peaks [32,44,45]. Numerous modified injection interfaces have been developed resulting in ${}^1W_b$ ranging from 3 to 500 ms [32,42,44–48]. An analyst should be able to produce ^{1}D peaks with ${}^1W_b$ ~ 100 ms, enabling a P_M of 50 ms to be used with ρ_S of ~ 2. A ${}^1n_c$ ~ 4000 (for a separation run time of 400 s) and a ${}^2n_c$ ~ 4 should be achievable resulting in a remarkable $n_{c,2D}$ (${}^1n_c$ ~ 4000) × (${}^2n_c$ ~ 4) ~16,000 or a peak capacity production of ~ 2400 peaks/min. Given this large peak capacity production, it is appealing to consider exploring comprehensive, ultra-fast GC × GC and GC^3.

An important consideration for any analytical technique is the ability to perform quantitative analysis. For a more rigorous test, chromatographic conditions (see Experimental) were purposely chosen to fully overlap cyclohexane and benzene on the ^{1}D separation. Fig. 2.8A shows the raw vector form of the data, and Fig. 2.8B shows the processed data, per Fig. 2.1. Following the data processing the ^{2}D peaks for benzene and cyclohexane are found to be baseline resolved (R_s greater than 1.5). Fig. 2.8C shows the processed data in the form of a GC × GC chromatogram. Because 100 μm i.d. columns were used, the apparent $^{2}W_b$ are extremely narrow with apparent $^{2}W_b$ of 8 and 9 ms for cyclohexane and benzene, respectively. Fig. 2.8D shows the calibration data for the different concentrations of cyclohexane and benzene. Peak areas were measured by summing all ^{2}D peaks above the limit of quantification (S/N of 10), i.e., along all ^{1}D modulations, to create a single ^{2}D peak [18]. Then, the single ^{2}D peak was integrated to provide a single peak area value. The measured peak areas for cyclohexane and benzene were both found to be linearly related to injected concentration, with R^2 values greater than 0.99 for each linear regression best fit line. As indicated by the standard deviation error bars, the quantitative precision in Fig. 2.8D was quite good, with an average RSD of 4.7% for analyte concentrations injected at or above 1.3 ppt.

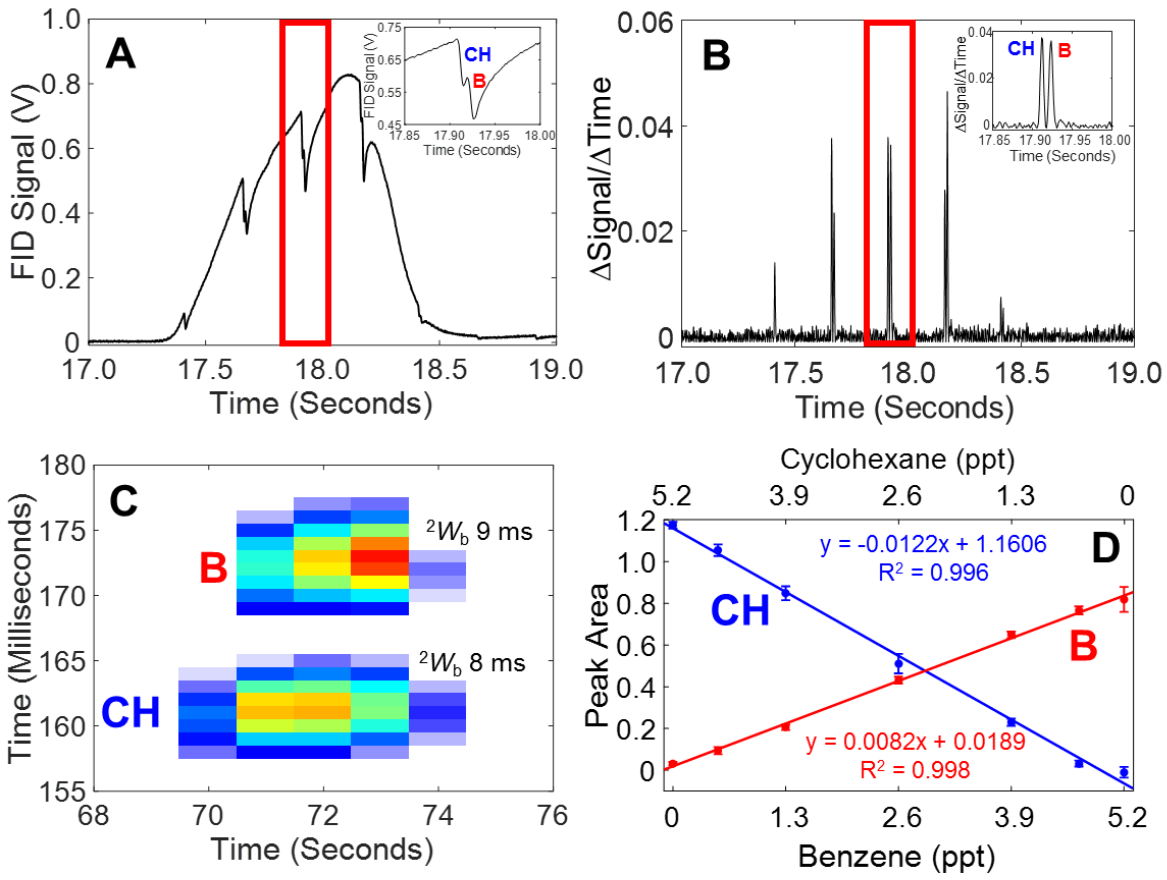


Figure 2.8 Partial modulation using the pulse flow valve was investigated to determine if the data is suitable to perform quantitative analysis, using mixtures of cyclohexane and benzene. (A) Raw vector form of the data for cyclohexane and benzene, labeled CH and B, respectively. (B) Processed data is shown, whereby the ^{2}D peaks are observed to be baseline resolved on the ^{2}D separation. (C) The processed data is shown in the form of a GC × GC separation with the apparent ^2W_b provided. (D) The concentration in ppt (parts per thousand) of cyclohexane and benzene were varied in order to study the quantitative aspects of the technique.

It is important to point out that the peak intensities of the processed data are directly related to the slope of the initial "error" function portion of the raw data. If an analyte is retained significantly on the ^{2}D separation, the "error" function in the raw data will have a smaller slope and following differentiation the "Gaussian" like ^{2}D peak will be wider with a concomitant lower intensity. If the "error" function becomes too broad, the analyte response may appear similar to the noise in the baseline. This issue is overcome as presented herein, by using highly efficient ^{2}D separations and operating in a low ^2k' separation regime. In addition to achieving acceptable precision for analyte quantification, reproducible ^2t_R were also achieved. The standard deviation for the ^2t_R for cyclohexane and benzene were both less than 1 ms indicating excellent reproducibility, limited by the FID collection rate of 1 kHz. This was determined by finding the retention times of all ^{2}D peaks above a *S/N* of 10 across all concentration ranges and injection replicates. There were about three ^{2}D peaks for each analyte per run, at six different concentrations (benzene and cyclohexane are not present at the opposite extremes of the calibration curves), and three injection replicates resulting in 54 measured ^2t_R.

2.4 CONCLUSION

A novel pulse flow valve modulator for GC × GC has been demonstrated, which creates narrow apparent $^2W_b \sim 10$ ms. A separation of a complex test mixture was demonstrated for various modulation periods P_M of 500, 250, 100 and 50 ms. A procedure has been introduced to convert the raw data into a format with the appearance of a GC × GC chromatogram through the process of differentiation. Relatively symmetric peaks were generated on both dimensions, with ^2t_R standard deviations less than 1 ms. In addition, the processed data has been shown to be suitable for quantitative analysis. Due to the extremely narrow peaks generated, the ^{2}D peak capacity is very high, resulting in a $^2n_c \sim 10$ for a modulation period of 250 ms. The

resulting $n_{c,2D}$ is ~ 4300 or a peak capacity production of 650 peaks/min which is 2-fold increase over traditional GC × GC applications. Furthermore, the pulse valve may be implemented in high speed, high peak capacity GC-sensors.

2.5 REFERENCES

[1] Z., Liu; J.B., Phillips, Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, J. Chromatogr. Sci. 29 (1991) 227-231.

[2] J., Dallüge; M., van Rijn; J., Beens; R.J.J., Vreuls; U.A.Th., Brinkman, Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection applied to the determination of pesticides in food extracts, J. Chromatogr. A 965 (2002) 207-211.

[3] J.M.D, Dimandja; S.B. Stanfil; J., Grainger; D.G., Patterson Jr., Application of comprehensive two-dimensional has chromatography (GC×GC) to the qualitative analysis of essential oils, J. High Resolut. Chromatogr. 23 (2000) 208-214.

[4] P, Korytar; P.E.G., Leonards; J., de Boer; U.A.Th., Brinkman, High-resolution separation of polychlorinated biphenyls by comprehensive two-dimensional gas chromatography, J. Chromatogr. A 958 (2002) 203-218.

[5] D., Ryan; R., Shellie; P., Tranchida; A., Casilli; L., Mondello; P. Marriott, Analysis of roasted coffee bean volatiles by using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, J. Chromatogr. A 1054 (2004) 57-65.

[6] H.J., de Geus; J., de Boer; U.A.Th. Brinkman, Development of a thermal desorption modulator for gas chromatography, J. Chromatogr. A 767 (1997) 137-151.

[7] B.V., Burger; T., Snyman; W.J.G., Burger; W.F., van Rooyen, Thermal modulator array for analyte modulation and comprehensive two-dimensional gas chromatography, J. Sep. Sci. 26 (2003) 123-128.

[8] J.B., Phillips; E.B., Ledford, Thermal modulation: A chemical instrumentation component of potential value in improving portability, Field Anal. Chem. Technol. 1 (1996) 23-29.

[9] H.J. de Geus; A., Schelvis; J., de Boer.; U.A.Th., Brinkman, Comprehensive two-dimensional gas chromatography with a rotating thermal desorption modulator and independently temperature-programmable columns, J. High Resolut. Chromatogr. 23 (2000) 189-196.

[10] P.J. Marriott, R.M. Kinghorn, Longitudinally modulated cryogenic system. A generally applicable approach to solute trapping and mobilization in gas chromatography, Anal. Chem. 69 (1997) 2582-2588.

[11] E.B. Ledford Jr., C. Billesbach, Jet-cooled thermal modulator for comprehensive multidimensional gas chromatography, J. High Resolut. Chromatogr. 23 (2000) 202-204.

[12] J. Beens, M. Adahchour, R.J.J. Vreuls, K. van Altena, U.A.Th. Brinkman, Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 919 (2001) 127-132.

[13] T., Hyötylăinen; M., Kallio; K., Hartonen; M., Jussila; S., Palonen; M.L., Riekkola, Modulator design for comprehensive two-dimensional gas chromatography: quantitative analysis of polyaromatic hydrocarbons and polychlorinated biphenyls, Anal. Chem. 74 (2002) 4441-4446.

[14] M., Pursch; P., Eckerle; J., Biel; R., Streck; H., Cortes; K., Sun; B., Winniford, Comprehensive two-dimensional gas chromatography using liquid nitrogen modulation: set-up and applications, J. Chromatogr. A 1019 (2003) 43-51.
[15] J., Harynuk; T., Gorecki, New liquid nitrogen cryogenic modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1019 (2003) 53-63.
[16] C.A., Bruckner; B.J., Prazen; R.E., Synovec, Comprehensive two-dimensional high-speed gas chromatography with chemometric analysis, Anal. Chem. 70 (1998) 2796-2804.
[17] J.V., Seeley; F., Kramp; C.J., Hicks, Comprehensive two-dimensional gas chromatography via differential flow modulation, Anal. Chem. 72 (2000) 4346-4352.
[18] C.E., Freye; L., Mu; R.E, Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1424 (2015) 127-133.
[19] A.E., Sinha; K.J., Johnson; B.J., Prazen; S.V., Lucas; C.G., Fraga; R.E., Synovec, Comprehensive two-dimensional gas chromatography of volatile and semi-volatile components using a diaphragm valve-based instrument, J. Chromatogr. A 983 (2003) 195-204.
[20] A.E., Sinha; B. J., Prazen; C.G., Fraga; R.E., Synovec, Valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection: instrumentation and figures-of-merit, J. Chromatogr. A 1019 (2003) 79-87.
[21] C.E., Freye; R.E., Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, Talanta 161 (2016) 675-680.
[22] D.R., Deans. A new technique for heart cutting in gas chromatography, Chromatographia 1 (1968) 18-22.
[23] J.V., Seeley; N.J., Micyus; S.V., Bandurski; S.K., Seeley; J.D., McCurry, Microfluidic Deans switch for comprehensive two-dimensional gas chromatography, Anal. Chem. 79 (2007) 1840-1847.
[24] P.A., Bueno; J.V., Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1027 (2004) 3-10.
[25] J.V., Seeley, Recent advances in flow-controlled multidimensional gas chromatography, J. Chromatogr. A 1255 (2012) 24-37.
[26] M., Edwards; A., Mostafa; T., Górecki, Modulation in comprehensive two-dimensional gas chromatography: 20 years of innovation, Anal. Bioanal. Chem. 401 (2011) 2335-2349.
[27] C., Duhamel; P., Cardinael; V., Peulon-Agasse; R., Firor; L., Pascaud; G., Semard-Jousset; P., Giusti; V., Livadaris, Comparison of cryogenic and differential flow (forward and reverse fill/flush) modulators and applications to the analysis of heavy petroleum cuts by high-temperature comprehensive gas chromatography, J. Chromatogr. A 1387 (2015) 95-103.
[28] J.F., Griffith; W.L., Winniford; K., Sun; R., Edam; J.C., Luong, A reversed-flow differential flow modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1226 (2012) 116-123.
[29] J.V., Seeley; S.K., Seeley, Comprehensive two-dimensional gas chromatography with pattern modulation, J. Chromatogr. A 1421 (2015) 114-122.
[30] H., Cai; S.D., Stearns, Partial modulation method via pulsed flow modulator for comprehensive two-dimensional gas chromatography, Anal. Chem. 76 (2004) 6064-6076.

[31] R.B., Wilson; W.C., Siegler; J.C., Hoggard; B.D., Fitz; J.S., Nadeau; R.E., Synovec, Achieving high peak capacity production for gas chromatography and comprehensive two-dimensional gas chromatography by minimizing off-column peak broadening, J. Chromatogr. A 1218 (2011) 3130-3139.
[32] J.L., Hope; K.J., Johnson; M.A., Cavelti; B.J., Prazen; J.W., Grate; R.E., Synovec, High-speed gas chromatographic separations with diaphragm valve-based injection and chemometric analysis as a gas chromatographic sensor, Anal. Chem. Acta. 490 (2003) 223-230.
[33] N.E., Watson; W.C., Siegler; J.C., Hoggard; R.E., Synovec, Comprehensive three-dimensional gas chromatography with parallel factor analysis, Anal. Chem. 79 (2007) 8270-8280.
[34] W.C., Siegler; J.A., Crank; D.W., Armstrong; R.E., Synovec, Increasing selectivity in comprehensive three-dimensional gas chromatography via an ionic liquid stationary phase column in one dimension, J. Chromatogr. A 1217 (2010) 3144-3149.
[35] N.E., Watson; H.D., Bahaghighat; K., Cui; R.E. Synovec, Comprehensive three-dimensional gas chromatography with time-of-flight mass spectrometry, Anal. Chem. 89 (2017) 1793-1800.
[36] B., Kehimkar; J.C., Hoggard; L.C., Marney; M.C., Billingsley; C.G., Fraga; T.J., Bruno; R.E., Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, J. Chromatogr. A 1327 (2014) 132-140.
[37] R.E., Murphy; M.R., Schure; J.P., Foley, Effect of sampling rate on resolution in comprehensive two-dimensional liquid chromatography, Anal. Chem. 70 (1998) 1585-1594.
[38] W., Khummueng; J., Harynuk; P.J., Marriott, Modulation ratio in comprehensive two-dimensional gas chromatography, Anal. Chem. 78 (2006) 4578-4587.
[39] W.C., Siegler; B.D., Fitz; J.C., Hoggard; R.E., Synovec, Experimental study of the quantitative precision for valve-based comprehensive two-dimensional gas chromatography, Anal. Chem. 83 (2011) 5190-5196.
[40] L.M., Blumberg, Accumulating resampling (modulation) in comprehensive two-dimensional capillary GC (GC×GC), J. Sep. Sci. 31 (2008) 3358-3365.
[41] R.E., Mohler; K.M., Dombek; J.C., Hoggard; E.T., Young; R.E., Synovec, Comprehensive two-dimensional gas chromatography time-of- flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells, Anal. Chem. 78 (2006) 2700-2709.
[42] R.B., Wilson; J.C., Hoggard; R.E., Synovec, High throughput analysis of atmospheric volatile organic compounds by thermal injection – isothermal gas chromatography – time-of-flight mass spectrometry, Talanta 103 (2013) 95-102.
[43] D.K., Pinkerton; B.A., Parsons; T.J., Anderson; R.E., Synovec, Trilinearity deviation ratio: A new metric for chemometric analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Anal. Chim. Acta. 871 (2015) 66-76.
[44] G.M., Gross; B.J., Prazen; J.W., Grate; R.E., Synovec, High-speed gas chromatography using synchronized dual-valve injection, Anal. Chem. 76 (2004) 3517-3524.

[45] H., Smith; R.D., Sacks, Column selectivity programming and fast temperature programming for high-speed GC analysis of purgeable organic compounds, Anal. Chem. 70 (1998) 4960-4966.
[46] T.-Y., Chen; M.-J., Li; J.-L., Wang, Sub-second thermal desorption of a micro-sorbent trap for the analysis of ambient volatile organic compounds, J. Chromatogr. A 976 (2002) 39-45.
[47] R.B., Wilson; B.D., Fitz; B.C., Mannion; T., Lai; R.K., Olund; J.C., Hoggard; R.E., Synovec, High-speed cryo-focusing injection for gas chromatography: reduction of injection band broadening with concentration enrichment, J. Chromatogr. A 97 (2012) 9-15.
[48] M.M., van Deursen; J., Beens; H.-G., Janssen; P.A., Leclercq; C.A., Cramers, Evaluation of time-of-flight mass spectrometric detection for fast gas chromatography, J. Chromatogr. A 878 (2000) 205-213.

Chapter 3. High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography

This chapter was reproduced from C.E. Freye, L. Mu, R.E. Synovec, "High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography" *Journal of Chromatography A* 1424 (2015) 127–133.

3.1 INTRODUCTION

One-dimensional gas chromatography (1D-GC) is an efficient analytical technique capable of separating volatile and semi-volatile species with high resolution. However, complex samples cannot always be sufficiently separated. Comprehensively adding a second dimension of GC increases the resolving power, and generally provides ample resolution for complex samples. This concept, theoretically pioneered by Giddings [1], was experimentally pioneered in the GC field by Phillips and Liu in 1991, who reported the first experiment for comprehensive two-dimensional gas chromatography (GC × GC) [2]. Use of GC × GC has been extremely beneficial for the analysis of many complex sample types such as such as fuels [3–5], environmental samples [6,7], pesticide contaminated foods [8], and fragrances [9–11]. In GC × GC, two sufficiently "orthogonal" columns are used to maximize the utilization of the two-dimensional separation space [12]. In order to transfer the analytes from the primary ^{1}D column to the secondary ^{2}D column in GC × GC, a modulation interface must be present. The GC × GC modulator serves as the heart of the instrument. The modulator operates by injecting portions of the ^{1}D column effluent onto the ^{2}D column with a user defined, relatively rapid frequency. There have been many developments and iterations in modulator design, and several reviews have been written on the various types of modulators applied with GC × GC [13–18]. There are basically three modulator

designs: thermal modulation [19–23], flow modulation [24–26], and diaphragm valve-based modulation [27–30], although the flow modulation and diaphragm valve modulation are often grouped together.

Thermal modulation is the most commonly applied, and relies on low temperatures to trap and focuses analytes as they elute from the ^{1}D column and introduces them to the ^{2}D column through rapid heating. The initial design pioneered by Phillips and Liu used a slotted heater which swept across a thick-filmed capillary which connected the ^{1}D column to the ^{2}D column [19,20,31]. The analytes eluting from the ^{1}D column were retained in the thick film and then volatilized and injected onto the ^{2}D column when the heater swept over the capillary. The most common type of thermal modulation is cryogenic modulation. In one type of cryogenic modulator design, developed by Marriott and Kinghorn, a jet of fluid CO_2 is applied to produce a moving cryogenic trap with a small portion of capillary between the two columns [21,32]. This causes the analytes to condense on the capillary and then is injected onto the ^{2}D column when heated rapidly. The most commonly applied cryogenic modulator design, developed by Ledford, Beens, and others, uses strategically placed and timed jets of cryogenic gas or a combination of heat and cooled jets to eliminate moving parts [22,33]. This design has been commercially adapted. While thermal modulation is very popular, it can be expensive to implement and is generally not amenable to reducing to a simple, lowest cost instrumental platform.

For the sake of discussion, modulators that employ flow switching and valve-based devices are grouped into the category of flow modulation. Dean switching has been used often in heart-cutting GC–GC to modulate sections of the effluent from the first column onto the second column [34]. Seeley first introduced the Deans switch as a modulator for GC × GC [24,34]. Agilent has introduced a capillary flow version (CFT) of the in-line fluidic modulator for

GC × GC, initially pioneered by Seeley, thus providing a commercially available way to modulate via a flow-based regime [35]. This commercial adaptation of the Deans switch-based GC × GC provides reproducible retention times on both dimensions a consistent performance, and hence, has been popular for industrial applications. Additionally, a differential flow modulator has been developed, which closely resembles a Deans switch but results in total transfer of the ^{1}D effluent [26]. Flow modulation, as categorized herein, is gaining popularity for GC × GC due to the relatively simple design and hearty performance [36,37].

The first diaphragm valve-based modulator design with GC × GC was reported in 1998, in which a diaphragm valve was programmed to divert ^{1}D column effluent to the ^{2}D column [27]. The original diaphragm valve-based GC × GC was limited to a maximum temperature of 175 °C, (due to shortcomings at that time in the valve construction for GC applications), and in the initial report only 10% of the ^{1}D effluent reached the ^{2}D column. However, diaphragm valve-based modulation is relatively simple and very inexpensive to implement, and use of a narrow injected pulse width onto the ^{2}D column has been demonstrated to produce a high GC × GC peak capacity even when a short modulation period is used (e.g., P_M of 1 to 2 s) [4,27,28,38,39]. Additional improvements to implementing a diaphragm valve-based modulator for GC × GC were reported by Seeley and co-workers, who employed a sample loop to collect the ^{1}D effluent, which provides a high fraction of the ^{1}D column effluent being transferred to the ^{2}D column [28]. When a high carrier flow rate on the ^{2}D column is used, sample pulses are compressed within the sample loop creating narrow ^{2}D peaks while increasing detection sensitivity. Indeed, using a diaphragm valve-based modulator, 100% transfer of the ^{1}D effluent to the ^{2}D separation has been reported, producing highly efficient separations with good peak capacity [30]. The primary shortcoming of the initial diaphragm valve-based modulator design in the past has been the limited temperature

range due to the O-rings within the valve, limiting the usage to a maximum of 175 °C if placed directly in the GC oven. This shortcoming was partially overcome by face mounting the valve on the top of the GC that extended the temperature limits to 265 °C while preserving the temperature sensitive O-rings within the valve [29]. By face mounting the valve, the temperature sensitive O-rings were not subjected to high temperatures, while the portions of the valve involved with sampling were within 1.2% of the oven temperature.

Recently, a state-of-the-art diaphragm valve with perfluoroelastomer O-rings has been developed that is capable of being placed directly in the GC oven and can readily withstand temperatures up to 325 °C for long term use. With this new valve, it is intriguing to consider a valve-based GC × GC for routine, low-cost, high efficiency separations. Indeed, herein we report a diaphragm valve-based GC × GC configuration that allows the valve to be used in temperatures up to 325 °C with the valve simply mounted directly in the oven. By operating the whole valve directly in the oven there is no need for complicated face mounting, and the whole valve is subject to the same temperature as the oven so there is no temperature difference between the oven and sample loop. This configuration will expand the ability of diaphragm valve-based GC × GC to analyze compounds with higher boiling points. A separation of a test mixture of alkanes, alcohols, and polyaromatic hydrocarbons with boiling points ranging from 98 to 450 °C is demonstrated. Retention time and peak width precision on the ^{2}D separations, and overall quantitative precision is evaluated. Additionally, valve-based GC × GC is compared to 1D-GC to determine the transfer fraction of the valve, as well as the relative detection sensitivity of the two techniques. Finally, we demonstrate applicability of this configuration to complex samples such as vacuum pump oil and orange oil.

3.2 EXPERIMENTAL

The GC × GC instrument with the high temperature diaphragm valve modulator was evaluated using a flame ionization detector (FID). The instrumental platform consisted of an Agilent 6890 GC (Agilent Technologies, Palo Alto, CA, USA) modified in-house with a high-speed, six-port diaphragm valve VICI model DV-12-1116 T (Valco Instruments Company Inc., Houston, TX, USA) fitted with a 5 μl sample loop which was mounted inside the oven. A diagram for the valve connections has been previously reported (albeit prior valve versions had temperature limitations), indicating the modulator connections, and direction of gas flows during the two modulation steps [14,28,29]. The diaphragm valve O-rings have been replaced with Kalrez O-rings allowing routine separation temperatures up to 325 °C.

The diaphragm valve was operated using LabVIEW 8.2 (National Instruments, Austin, TX, USA) program which was used to control the valve actuation. The stock electrometer for the Agilent FID was replaced with a high-speed electrometer built in-house allowing the data to be collected at 100 kHz, with the data further boxcar averaged to 1 kHz followed by Savitsky–Golay smoothing. Collecting the FID signal at a higher frequency than the stock electrometer facilitates more precise data smoothing operations in order to optimize the signal-noise-ratio (*S/N*) concurrent with not introducing additional broadening of the ^{2}D peaks.

Samples were introduced to the GC × GC via a 7683B auto-injector (Agilent Technologies, Palo Alto, CA, USA). Data was collected at 1 kHz with an FID temperature of 325 °C. The ^{1}D column was a RTX-5ms (diphenyl dimethyl polysiloxane) stationary phase: 20 m length, 250 μm inner diameter (i.d.), and 0.25 μm film thickness. The ^{2}D column was a RTX-200 (trifluoropropylmethyl polysiloxane) stationary phase: 2.4 m length, 180 μm i.d., and 0.2 μm film thickness. Prior to each sample injection, hexane and acetone were used as solvent rinses. The inlet

temperature was set to 325 °C. The flow on ^{1}D was 1.5 ml/min while flow on ^{2}D was 20 ml/min. Oven temperature control on ^{2}D was not applied, so both columns were at the same nominal temperature. Column separation conditions were selected in order to sufficiently use the two dimensional separation space for the test mixture, the vacuum pump oil, and the orange oil, with the following specific conditions for each sample.

A 2 μl volume of the test mixture which contained alkanes, alcohols, and polyaromatic hydrocarbons was injected at a split of 20:1. For the test mixture separation, the oven was held at 40 °C for 3 min and ramped at 10 °C/min to 325 °C where it was held for 1 min, and the modulation period P_M was 1.0 s with an injection pulse width of 25 ms. Five replicates were obtained for the test mixture. For the analysis of orange oil, 1 μl was injected with a 20:1 split, the oven program was identical to the test mixture analysis except the oven was held at 325 °C for 3 min, and the modulation period P_M was 1.5 s with an injection pulse width of 25 ms. The vacuum pump oil was diluted 1:2 in hexane. A 2 μl sample volume of diluted vacuum pump oil was injected with a 100:1 split, the oven was held at 200 °C for 1 min and ramped at 5 °C/min to 325 °C and held for 1 min, and the modulation period P_M was 1.0 s with an injection 25 ms pulse width. The vacuum pump oil was also run on a GC × GC–TOFMS with similar separation conditions to aid in identification of selected unresolved analytes. The GC × GC–TOFMS instrumental platform consisted of an Agilent 6890 N gas chromatograph equipped with an Agilent 7683 autoinjector (Agilent Technologies, Palo Alto, CA) coupled with a LECO Pegasus III TOFMS equipped with a 4D thermal modulator upgrade (LECO, St. Joseph, MI). The electron energy for the TOFMS was set to −70 eV with a detector voltage of 1600 V and data was collected from 35 amu to 334 amu at unit mass resolution.

3.3 RESULTS AND DISCUSSION

The GC × GC–FID chromatogram for the test mixture (Table 3.1) is shown in Fig. 3.1A. The chromatogram has been re-registered on the ^{2}D dimension to remove the dead time. On the reregistered ^{2}D time axis, the alkanes elute at ~0.3 s, the alcohols elute at ~0.4 s and the polyaromatic hydrocarbons elute from ~0.5 to 0.8 s. The earliest analyte, heptane is eluting at 2.2 min and an oven temperature of 40 °C while the latest eluting peak, triacontane, elutes at 26.6 min which corresponds to an oven temperature of 276 °C. A comprehensive list of elution times is provided in Table 3.1.

Test Mixture							
Compound	**Boiling Point °C**	**$^{1}t_{R}$ (min)**	**$^{2}t_{R}$ (s)**	**%RSD**	**$^{2}W_{b}$ (ms)**	**%RSD**	**%RSD Peak Area**
Alkanes							
Heptane	98	2.2	0.5	0.4	67.2	3.6	8.1
Octane	125	3.5	0.5	0.2	72.7	5.7	*
Nonane	151	5.2	0.5	0.2	73.4	3.5	2.0
Decane	174	6.8	0.6	0.1	66.2	6.7	1.8
Undecane	196	8.4	0.6	0.3	78.3	0.9	1.9
Dodecane	216	9.8	0.6	0.3	75.5	2.1	4.1
Tridecane	235	11.2	0.6	0.6	77.5	3.3	4.2
Tetradecane	254	12.5	0.6	0.4	76.5	2.0	3.6
Pentadecane	271	13.7	0.6	0.4	75.1	3.1	4.0
Hexadecane	287	14.9	0.6	0.3	76.3	4.1	4.1
Heptadecane	302	15.9	0.6	0.3	78.9	2.3	4.3
Octadecane	317	17.0	0.6	0.3	77.8	3.0	3.4
Nonadecane	330	18.0	0.6	0.3	75.6	4.1	4.6
Eicosane	343	18.9	0.6	0.2	76.9	2.2	1.3
Heneicosane	357	19.8	0.6	0.3	78.9	3.1	1.4
Docosane	368	20.7	0.6	0.1	76.3	1.6	2.0
Tricosane	380	21.5	0.6	0.3	75.5	3.6	3.1
Tetracosane	391	22.3	0.6	0.4	77.0	3.6	3.3
Pentacosane	401	23.1	0.6	0.4	76.3	3.2	3.0
Hexacosane	412	23.9	0.6	0.5	77.7	2.3	2.3
Heptacosane	422	24.6	0.6	0.4	77.4	2.8	5.5
Octacosane	432	25.3	0.6	0.4	76.3	2.9	4.4
Nonacosane	441	26.0	0.6	0.5	77.3	3.3	6.7
Triacontane	450	26.6	0.6	0.1	75.1	1.2	7.8

Alcohols							
Hexanol	157	4.7	0.7	0.2	72.6	2.6	7.1
2-Heptanol	159	5.2	0.7	0.2	76.0	2.4	5.3
Octanol	195	8.0	0.7	0.3	77.3	3.0	8.5
Nonanol	215	9.4	0.7	0.7	76.8	1.4	6.6
Decanol	231	10.8	0.7	0.6	79.1	3.2	4.2
Dodecanol	256	13.4	0.7	0.4	80.3	3.7	8.1
Tetradecanol	289	15.7	0.7	0.3	84.4	3.1	4.5
Hexadecanol	344	17.8	0.7	0.4	80.8	2.8	5.3
Octadecanol	210 at 15 mm Hg	19.7	0.7	0.3	83.1	3.3	2.6
Eicosanol	370	21.5	0.7	0.3	85.9	6.5	7.8
Docosanol	180 at 0.22 mm Hg	23.1	0.7	0.4	90.1	2.1	3.9
Tetracosanol	395	24.6	0.7	0.2	84.5	3.5	3.8
Hexacosanol	446	26.0	0.7	0.4	81.8	5.5	7.7
Polyaromatic hydrocarbons							
Naphthalene	218	9.7	0.8	0.13	78.5	3.8	3.1
Fluorene	295	14.8	0.9	0.2	83.0	1.9	4.2
Phenanthrene	332	16.9	0.9	0.2	95.8	1.5	4.9
Anthracene	340	17.0	0.9	0.2	88.7	3.4	1.5
Fluoranthene	375	19.7	0.04	4.1	95.9	3.4	3.5
Pyrene	404	20.1	0.04	2.6	93.4	3.0	1.1
Chrysene	448	23.0	0.1	3.9	91.3	2.3	9.8

Table 3.1 List of the analytes in the test mixture as well as their boiling points, retention times on both dimensions, width at base ^2W_b on the ^{2}D dimension, and the %RSD for ^2W_b. Additionally, peak area of the summed ^{2}D peak was quantified and the %RSD is reported ([a]Octane was used as an internal standard).

In order to discuss the merits of the high temperature diaphragm valve-based GC × GC, an enhanced view of a portion of the separation of the text mixture including tetradecanol is shown in Fig. 3.1B (box indicated in Fig. 3.1A). Tetradecanol, a representative mid-separation eluting alcohol in the mixture, has an apparent width on the first dimension ^1W_b of 2.5 s which was determined by fitting a Gaussian function over the data (through the apexes of each of the ^{2}D peaks in the raw vector format) [40–42]. With a modulation period P_M of 1.0 s, the modulation ratio M_R which is defined as $^1W_b/P_M$ was determined to be 2.5 for tetradecanol [41].

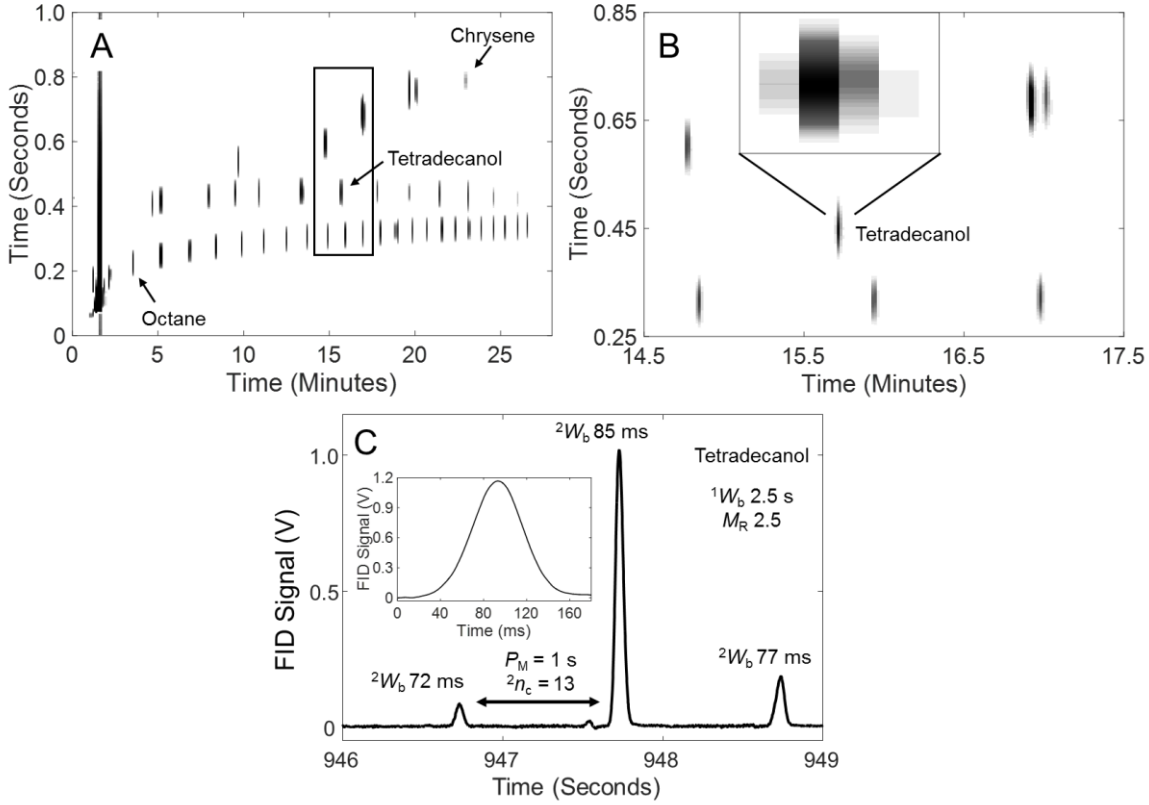


Figure 3.1 (A) GC × GC – FID chromatogram of the test mixture separation which has been registered to adjust the ^{2}D dead time. Analytes span an elution temperature range of 40–276 °C. On the reregistered ^{2}D time axis, alkanes elute at ~0.3 s, alcohols elute at ~0.4 s, and polyaromatic hydrocarbons elute from 0.5 to 0.8 s. A comprehensive peak list with retention times and peak widths is provided in Table 3.1. (B) Enhanced view of a selected portion of the chromatogram containing tetradecanol is shown. (C) Shown is the raw vector form of the GC × GC – FID data for tetradecanol. Peaks in the ^{2}D separation are narrow yielding significant peak capacity on the ^{2}D dimension even in conjunction with a short modulation period P_M of 1.0 s. The insert Fig. is the sum of all ^{2}D peaks to produce a single ^{2}D peak.

Additional figures-of-merit are discussed using the raw chromatogram of tetradecanol (in Fig. 3.1C), which elutes at 15.7 min, corresponding to an elution temperature of 167 °C. The average ^2W_b was 78.0 ms, corresponding to a peak capacity of $^2n_c \sim 13$ at unit resolution, which is very acceptable for the modulation period P_M of 1.0 s [28,29,42]. Peak capacity on ^{2}D was calculated at unit resolution with $R_S = 1$, defined by the P_M divided by the average ^2W_b. On an average, the alcohols exhibited very low variation in retention time on ^{2}D dimension with a %RSD of 0.4%, and they had an average ^2W_b of 81.0 ms with an average %RSD of 3.3%. Regarding our overall observations, when the valve is switched from the fill to flush position, the volumetric flow rate of the mobile phase decreases temporarily to bring the sample loop up to proper ^{2}D column pressure which appears to result in minor flow fluctuations. However, for analytes eluting near the actuations no noticeable deformations in ^{2}D peak shapes were observed, albeit with some minor ^{2}D retention time shifting. No tailing was observed on ^{2}D, despite having some very polar, low volatility compounds in the test mixture. This implies the wetted surfaces of the valve are behaving very inert with minimal dead volume.

The quantitative area for the GC × GC peaks was also determined by summing all ^{2}D peaks above the limit of quantification (LOQ), defined by a *S/N* of 10. The insert for Fig. 3.1C shows the summation of all the ^{2}D peaks for tetradecanol. The summed ^{2}D peak was then integrated to determine the quantitative peak area for tetradecanol. The %RSD for the calculation of the peak area for tetradecanol was 4.5% (for the five trials), and at a M_R of 2.5, this %RSD is consistent with previous reports [42]. The peak areas were calculated for all 44 analytes in the test mixture. A comprehensive list of ^{2}D peak widths and their respective width %RSD is provided in Table 3.1 as well as the %RSD for their quantification using octane as the internal standard (see Fig. 3.1A). The average %RSD for the peak area for all compounds was 4.4%.

We now turn our attention to the peak capacity provided by the instrument. Using the data for all 44 analytes in the test mixture (Table 3.1), based on the average peak width on the ^{1}D dimension of 2.4 s with a run time of 32.5 min, the ^{1}D peak capacity is $^1n_c \sim 800$. The average peak width on the ^{2}D dimension is 79.4 ms (see Table 3.1) with a modulation period P_M of 1.0 s results in a ^{2}D peak capacity of $^2n_c \sim 12$. Thus, the ideal two-dimensional peak capacity given by $n_{c,2D} = {}^1n_c \times {}^2n_c$ is 9600 or a peak capacity production of about 300 peaks/min. The two-dimensional peak capacity of this separation is very high, and is similar to other reports using valve-based GC × GC [28,42]. Overall, the peaks are quite narrow on the ^{2}D dimension leading to a high two-dimensional peak capacity even with a short modulation period P_M of 1.0 s, which was necessary due to the very narrow peaks produced on the ^{1}D separation.

The valve-based GC × GC was compared to 1D-GC to determine the transfer fraction of the valve, as well as the relative sensitivity of the two techniques. The 1D-GC chromatogram is shown in Fig. 3.2A. The same temperature program and flow rate were used with 1D-GC in order to obtain a separation essentially equivalent to the ^{1}D separation obtained with GC × GC (Fig. 3.1A). Near the end of the separation, one can visibly see the baseline drift due to the column bleed at higher temperature. The valve is readily operational up to 325 °C while typical GC columns have a temperature limit around 325 °C, while polar columns (such as polyethylene glycol (Rtx-Wax) or cyanopropylmethyl phenylmethyl polysiloxane (Rtx-225)) often have temperature limits of roughly 250 °C). Two representative analytes, tridecane, and triacontane, were used in this instrumentation comparison. Fig. 3.2B shows the 1D-GC peak for tridecane. The raw vector form for the GC × GC data has been overlaid on the 1D-GC peak for comparison. The modulated

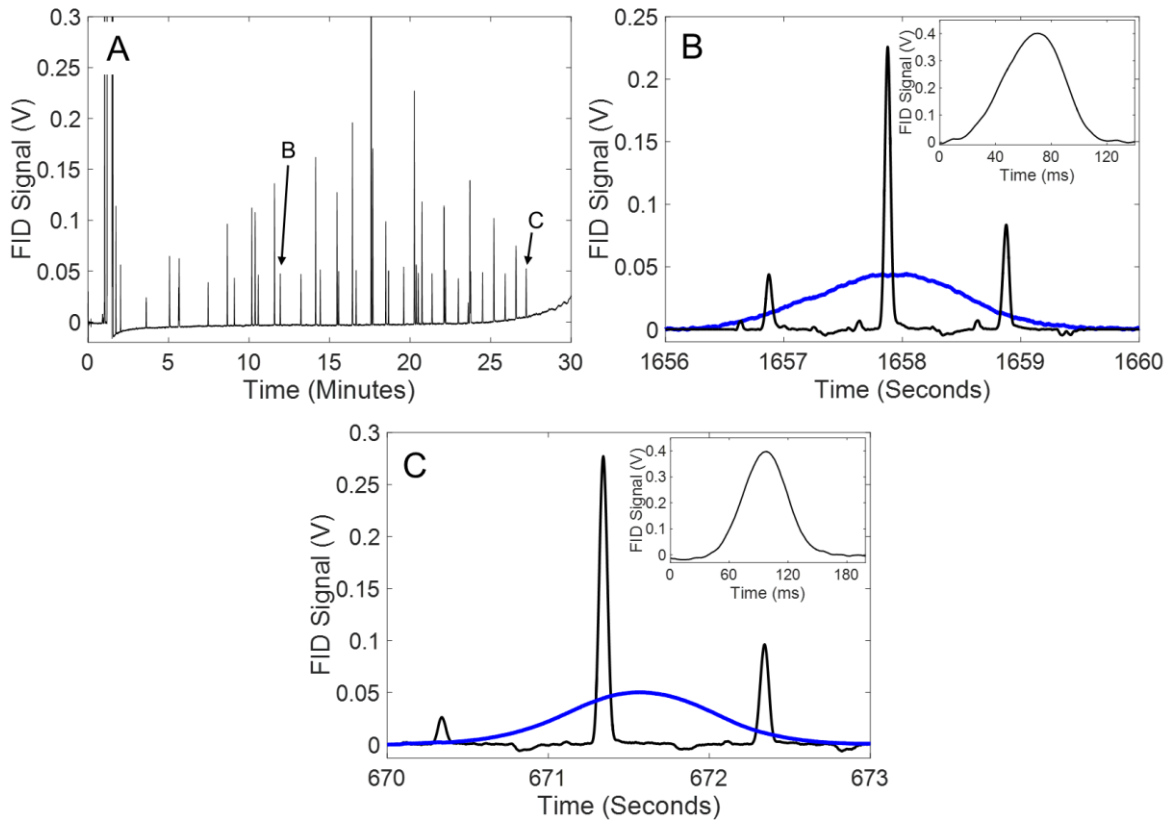


Figure 3.2 (A) GC–FID chromatogram of the test mixture is shown. (B) and (C) refer to tridecane and triacontane which are used to discuss transfer fraction of the valve and the sensitivity of GC × GC relative to 1D-GC. (B) Shown is the overlaid 1D-GC and GC × GC data for tridecane. The insert shows the summed ^{2}D peaks to give a single tridecane ^{2}D peak for the GC × GC data. (C) Shown is the overlaid 1D-GC and GC × GC data for triacontane. The insert shows the summed ^{2}D peaks to give a single triacontane ^{2}D peak for the GC × GC data.

GC × GC peaks on ^{2}D elute 1 s later than shown (since P_M = 1 s), but are shown directly on top of the 1D-GC peak for clarity. With valve modulation, the transfer fraction of tridecane from ^{1}D to ^{2}D was determined to be 36% at a M_R of 2.0. The transfer fraction was determined by integrating the 1D-GC and comparing that peak area to the integrated GC × GC peak area. The insert for Fig. 3.2B shows the summed ^{2}D peaks to give a single tridecane ^{2}D peak for the GC × GC peak area

determination. In terms of detected peak sensitivity, although 36% of tridecane is transferred in GC × GC, the most intense ^{2}D peak is 5.5 times more intense than the 1D-GC peak. This enhanced detection sensitivity for GC × GC is a consequence of zone compression with valve-based GC × GC [26]. Indeed, the ^{2}D tridecane peaks is ~23 times narrower than the single 1D-GC peak. In fact, if one considers the summed ^{2}D peak in the insert Fig. 3.2B with a height of ~0.4 V, it is~8 times higher than the 1D-GC peak with a height of~0.05 V. Because the tridecane peaks are 23-fold narrower on ^{2}D for GC × GC than for the single 1D-GC peak, the 36% transfer fraction of the valve is more than offset, resulting in the 8-fold higher sensitivity using the summed ^{2}D peak in the comparison. Similarly, Fig. 3.2C shows the 1D-GC and GC × GC peaks for triacontane. The transfer fraction from ^{1}D to ^{2}D was 27% for triacontane at M_R of 2.3. The insert shows the summed ^{2}D peaks for triacontane. Similar to the tridecane, the most intense ^{2}D peak is ~5.5 times the intensity of the 1D-GC peak, however the 1D-GC peak is ~29 times the width of the most intense ^{2}D peak.

Although vacuum pump oil does not contain a large diversity of chemical compound classes, it is an excellent test mixture for showing the ability of the high temperature diaphragm valve-based GC × GC – FID instrument to perform at higher operating temperatures. Fig. 3.3A shows the separation of the vacuum pump oil. The first analyte elutes at an oven temperature of ~220 °C while the last analyte elutes at an oven temperature of ~315 °C. Because vacuum pump oil is a very refined hydrocarbon sample, it is not very complex in terms of having different classes of analytes. A one dimensional view (i.e., raw vector form) of a portion of the separation is provided in Fig. 3.3B. Peak widths for the vacuum pump oil on the ^{2}D dimension are significantly wider than those determined in the test mixture. When the sample was run on the GC × GC – TOFMS to determine peak identity, it was discovered that the peaks that initially appeared to be

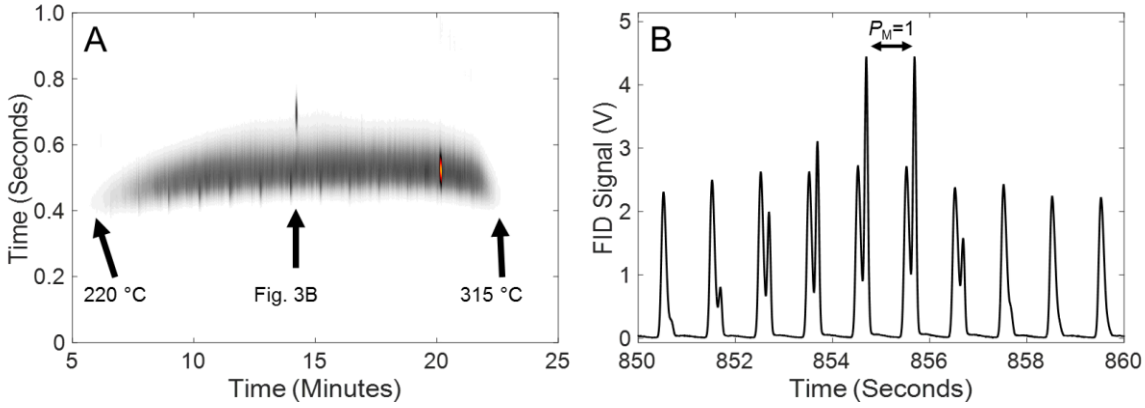


Figure 3.3 (A) The GC × GC–FID chromatogram for vacuum pump oil is shown. The elution temperature range for detected peaks spans from 220 °C to 315 °C. (B) A detailed view is shown of an analyte peak at a ^1t_R and ^2t_R of 14 min and 0.7 s, respectively.

single analytes with GC × GC – FID actually contained several overlapping analytes which in turn resulted in a slightly inflated peak widths in the GC × GC – FID separation. Based on retention times and mass spectral information, interpretation of the data suggested that the vacuum pump oil contains primarily cyclic compounds. Because there are several analytes contained within each peak, it is not appropriate to calculate peak widths. Nevertheless, based upon inspection of the peak shapes across the entire chromatogram (with those in Fig. 3.3B being representative), the peaks are reasonably symmetric, especially at higher temperatures.

Orange oil has commonly been studied using one dimensional gas chromatography–mass spectrometry due to the complex sample containing a wide range of compounds with an extremely wide range in volatility, and it is one of more prevalent essential oils. Hence, orange oil is ideally suited to challenge the performance of our high temperature diaphragm valve-based GC × GC – FID instrument. The chromatogram of orange oil is shown in Fig. 3.4A; we see that the orange oil sample is complex and GC × GC is necessary to achieve necessary resolution between a majority of the analytes. The first analyte elutes from the ^{1}D column at 1.0 min corresponding to an oven

temperature of 40 °C, while the last analyte elutes at 30.2 min corresponding to an oven temperature of 312 °C. In the chromatogram of the orange oil, the ^{1}D column starts to bleed at 30 min.

A portion of the orange oil separation in Fig. 3.4A is presented in Fig. 3.4B to facilitate discussion of the performance of the high temperature diaphragm valve-based GC × GC–FID instrument. The portion of the complete orange oil separation in Fig. 3.4B is from 25 to 31 min (an oven temperature of 260–320 °C), and displays the significant resolution obtained through narrow ^{2}D peak widths. However, raw vector form data provides further insight to discuss the merits of this separation. Fig. 3.4C and D show two "unknown" analytes from Fig. 3.4B in the raw vector form. These unknown analytes are highly retained analytes on the ^{1}D and ^{2}D dimensions with a reasonably high signal-to-noise ratio. The average $^{2}W_{b}$ for Analyte 1 in Fig. 3.4C is 142.0 ms, corresponding to a peak capacity of $^{2}n_{c} \sim 11$. Analyte 2 in Fig. 3.4D has an average $^{2}W_{b}$ of 141.0 ms, corresponding to a peak capacity of $^{2}n_{c} \sim 11$. Also, Analytes 1 and 2 have $^{1}W_{b}$ of 4.2 and 3.1 s which corresponds to a M_{R} of 2.8 and 2.1, respectively. Based on the average peak width on ^{1}D dimension (3.7 s) with a run time of 34.5 min, the ^{1}D peak capacity is $^{1}n_{c} \sim 560$. The average peak width on the ^{2}D dimension is 122.3 ms with a modulation period P_{M} of 1.5 s results in a ^{2}D peak capacity of $^{2}n_{c} \sim 12$. With a modulation ratio M_{R} of 2.8, the ideal two-dimensional peak capacity is $n_{c,2D} \sim 6000$ or a peak capacity production of roughly 175 peaks/min spanning an extensive temperature range of 285 °C. As seen in Fig. 3.4C and D, Analytes 1 and 2 both have symmetrical peak shapes on the ^{2}D dimension. At higher temperatures, thermal modulators often tail because the analytes may not desorb uniformly due to their high boiling points. Based upon this initial study with this newly available valve, we surmise that

diaphragm valve-based GC × GC is now limited by the column chemistry and no longer by the valve temperature limits.

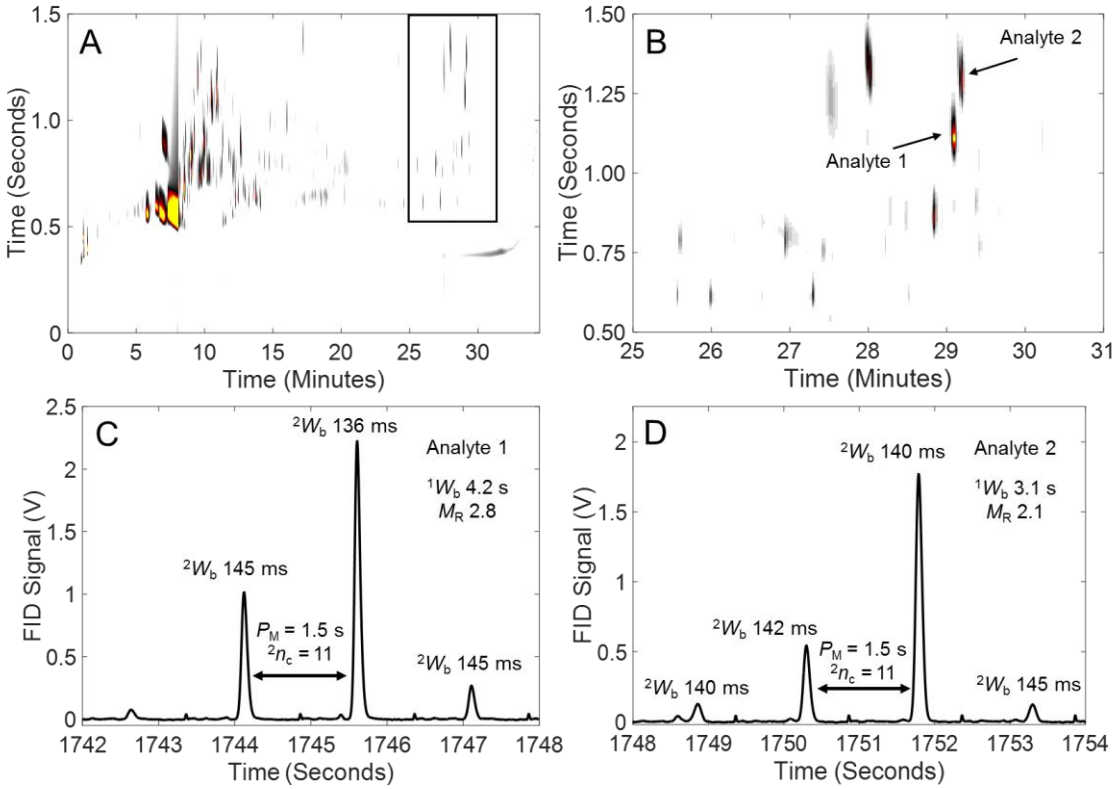


Figure 3.4 (A) Shown is the GC × GC – FID chromatogram for orange oil. (B) A detailed portion of the chromatogram is shown from 25 to 31 min which corresponds to temperatures of 260–320 °C. (C) A detailed view is shown of an Analyte 1 eluting at 29.1 min. Due to the narrow peaks, peak capacity is significant on the ^{2}D dimension and the peaks are uniform in shape even at high temperatures. (D) A detailed view is shown of the Analyte 2 eluting at 29.2 min.

3.4 CONCLUSION

A high temperature diaphragm valve-based GC × GC – FID has been developed and initially studied. By using a recently developed diaphragm valve, in which the stock O-rings (limited to 175 °C) have been replaced with Kalrez O-rings, the temperature limit has been extended to 325 °C with the valve mounted directly in the oven. Separation of the test mixture of alkanes, alcohols, and polyaromatic hydrocarbons which encompasses a wide range of boiling points produced uniform, symmetric peaks on the both the ^{1}D and ^{2}D dimensions indicating inertness of the wetted portions of the valve. High reproducibility was observed for ^{2}D peak widths, retention times, and peak areas. Due to narrow peaks on the ^{2}D dimension, the peak capacity is very useful with $^{2}n_{c}$ ranging from 11 to 13, even with a short modulation period, which facilities producing truly comprehensive GC × GC separation [29,38,42] with an adequate modulation ratio [41]. Separations of complex mixtures such as vacuum pump oil and orange oil demonstrated extended applicability of diaphragm valve-based GC × GC to 325 °C, which is roughly the temperature limit of commonly used GC columns.

3.5 ACKNOWLEDGEMENTS

We would like to thank Stan Stearns of Valco for providing the high temperature valves.

3.6 REFERENCES

[1] J.C. Giddings, Two-dimensional separations: concept and promise, Anal. Chem. 56 (1984) 1258A-1270A.

[2] Z. Liu, J.B. Phillips, Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, J. Chromatogr. Sci. 29 (1991) 227-231.

[3] P.J. Schoenmakers, J.L.M.M. Oomen, J. Blomberg, W. Genuit, G. van Velzen, Comparison of comprehensive two-dimensional gas chromatography and gas chromatography – mass spectrometry for the characterization of complex hydrocarbon mixtures, J. Chromatogr. A 829 (2000) 29–46.

[4] C.G. Fraga, B.J. Prazen, R.E. Synovec, Comprehensive two-dimensional gas chromatography and chemometrics for the high-speed quantitative analysis of aromatic isomers in jet fuel using the standard addition method and an objective retention time alignment algorithm, Anal. Chem. 72 (2000) 4154-4162.

[5] K.J. Johnson, R.E. Synovec, Pattern recognition of jet fuels: comprehensive GC × GC with ANOVA-based feature selection and principal component analysis, J. Chemom. Intell. Lab. Syst. 60 (2002) 225–237.
[6] A.C. Lewis, N. Carslaw, P.J. Marriott, R.M. Kinghorn, P. Morrison, A.L. Lee, K.D. Bartle, M.J. Pilling, A larger pool of ozone-forming carbon compounds in urban atmospheres, Nature 404 (2002) 778–781.
[7] P. Korytar, P.E.G. Leonards, J. de Boer, U.A.Th. Brinkman, High-resolution separation of polychlorinated biphenyls by comprehensive two-dimensional gas chromatography, J. Chromatogr. A 958 (2002) 203–218.
[8] J. Dallüge, M. van Rijn, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection applied to the determination of pesticides in food extracts, J. Chromatogr. A 965 (2002) 207–211.
[9] R. Shellie, L. Mondello, P. Marriott, G. Dugo, Characterization of lavender essential oils by using gas chromatography–mass spectrometry with correlation of linear retention indices and comparison with comprehensive two-dimensional gas chromatography, J. Chromatogr. A 970 (2002) 225–234.
[10] J.M.D. Dimandja, S.B. Stanfill, J. Grainger, D.G. Patterson, Jr., Application of comprehensive two-dimensional gas chromatography (GC×GC) to the qualitative analysis of essential oils, J. High Resol. Chromatogr. 23 (2000) 208–214.
[11] P.J. Marriott, R. Shellie, C. Cornwell, Gas chromatographic technologies for the analysis of essential oils, J. Chromatogr. A 936 (2001) 1–22.
[12] Z. Liu, D.G. Patterson Jr., M.L. Lee, Geometric approach to factor analysis for the estimation of orthogonality and practical peak capacity in comprehensive two-dimensional separations, Anal. Chem. 67 (1995) 3840-3845.
[13] P. Marriott, R. Shellie, Principles and applications of comprehensive two-dimensional gas chromatography, Trends Anal. Chem. 21 (2002) 573-583.
[14] J.V. Seeley, Recent advances in flow-controlled multidimensional gas chromatography, J. Chromatogr. A 1255 (2012) 24– 37.
[15] J. Harynuk, T. Gorecki, New liquid nitrogen cryogenic modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr., A 1019 (2003) 53–63.
[16] P.J. Marriott, P.D. Morrison, R.A. Shellie, M.S. Dunn, E. Sari, D. Ryan, Recent applications in multidimensional chromatography, LCGC Eur. 16 (2003) 23-31.
[17] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC × GC): I. Introduction and instrumental set-up, Trends Anal. Chem. 25 (2006) 438-454.
[18] P.Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, Modulators for comprehensive two-dimensional gas chromatography, Trends Anal. Chem. 30 (2011) 1437-1461.
[19] J.B. Phillips, R.B. Gaines, J. Blomberg, F.W.M. van der Wielen, J.M. Dimandja, V. Green, J. Granger, D. Patterson, L. Racovalis, H.J. de Geus, J. de Boer, P. Haglund, J. Lipsky, V. Sinha, E.B. Ledford Jr., A robust thermal modulator for comprehensive two-dimensional gas chromatography, J. High Resolut. Chromatogr. 22 (1999) 3-10.
[20] J.B. Phillips, J. Xu, Comprehensive multi-dimensional gas chromatography, J. Chromatogr. A 703 (1995) 327-334.
[21] P.J. Marriott, R.M. Kinghorn, Longitudinally modulated cryogenic system. A generally applicable approach to solute trapping and mobilization in gas chromatography, Anal. Chem. 69 (1997) 2582-2588.

[22] J. Beens, M. Adahchour, R.J.J. Vreuls, K. van Altena, U.A.Th. Brinkman, Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 919 (2001) 127-132.
[23] J. Harynuk, T. J. Gorecki, Comprehensive two-dimensional gas chromatography in stop-flow mode, J. Sep. Sci. 27 (2004) 432-441.
[24] J.V. Seeley, N.J. Micyus, S.V. Bandurski, S.K. Seeley, J.D. McCurry, Microfluidic Deans switch for comprehensive two-dimensional gas chromatography, Anal. Chem. 79 (2007) 1840-1847.
[25] B.D. Quimby, J.D. McCurry, W.M. Norman, Capillary flow technology for gas chromatography: Reinvigorating a mature analytical discipline, LCGC The Peak (2007) 7-15.
[26] P.A. Bueno, J.V. Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1027 (2004) 3-10.
[27] C.A. Bruckner, B.J. Prazen, R.E. Synovec, Comprehensive two-dimensional high-speed gas chromatography with chemometric analysis, Anal. Chem. 70 (1998) 2796-2804.
[28] J.V. Seeley, F. Kramp, C.J. Hicks, Comprehensive two-dimensional gas chromatography via differential flow modulation, Anal. Chem. 72 (2000) 4346-4352.
[29] A.E. Sinha, K.J. Johnson, B.J. Prazen, S.V. Lucas, C.G. Fraga, R.E. Synovec, Comprehensive two-dimensional gas chromatography of volatile and semi-volatile components using a diaphragm valve-based instrument, J. Chromatogr. A 983 (2003) 195-204.
[30] R.E. Mohler, B.J. Prazen, R.E. Synovec, Total-transfer, valve-based comprehensive two-dimensional gas chromatography, Anal. Chim. Acta 555 (2006) 68-74.
[31] H.J. de Geus, A. Schelvis, J. de Boer, U.A.Th. Brinkman, Comprehensive two-dimensional gas chromatography with a rotating thermal desorption modulator and independently temperature-programmable columns, J. High Resolut. Chromatogr. 23 (2000) 189-196.
[32] R.M. Kinghorn, P.J. Marriott, Design and implementation of comprehensive gas chromatography with cryogenic modulation, J. High Resolut. Chromatogr. 23 (2000) 245-252.
[33] E.B. Ledford Jr., C. Billesbach, Jet-cooled thermal modulator for comprehensive multidimensional gas chromatography, J. High Resolut. Chromatogr. 23 (2000) 202-204.
[34] D.R. Deans, A new technique for heart cutting in gas chromatography, Chromatographia 1 (1968) 18-22.
[35] B.D. Quimby, J.D. McCurry, W.M. Norman, Capillary flow technology for gas chromatography: reinvigorating a mature analytical discipline, LCGC The Peak (2007) 7-15.
[36] M. Poliak, M. Kochman, A. Amirav, Pulsed flow modulation comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1186 (2008) 189-195.
[37] Q. Gu, F. David, F. Lynen, K. Rumpel, G. Xu, P. Vos, P. Sandra, Analysis of bacterial fatty acids by flow modulated comprehensive two-dimensional gas chromatography with parallel flame ionization detector/mass spectrometry, J. Chromatogr. A 1217 (2010) 4448-4453.
[38] B.J. Prazen, K.J. Johnson, A.W. Weber, R.E. Synovec, Two-dimensional gas chromatography and trilinear partial least squares for the quantitative analysis of aromatic and naphthene content in naphtha, Anal. Chem. 73 (2001) 5677-5682.

[39] J.V. Seeley, F.J. Kramp, K.S. Sharpe, S.K. Seeley, Characterization of gaseous mixtures of organic compounds with dual-secondary column comprehensive two-dimensional gas chromatography, J. Sep. Sci. 25 (2002) 53-59.
[40] R.E. Murphy, M.R. Schure, J.P. Foley, Effect of sampling rate on resolution in comprehensive two-dimensional liquid chromatography, Anal. Chem. 70 (1998) 1585-1594.
[41] W. Khummueng, J. Harynuk, P.J. Marriott, Modulation ratio in comprehensive two-dimensional gas chromatography, Anal. Chem. 78 (2006) 4578-4587.
[42] W.C. Siegler, B.D. Fitz, J.C. Hoggard, R.E. Synovec, Experimental study of the quantitative precision for valve-based comprehensive two-dimensional gas chromatography, Anal. Chem. 83 (2011) 5190-5196.

Chapter 4. High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry

This chapter was reproduced from C.E. Freye, R.E. Synovec, "High Temperature Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry" *Talanta* 161 (2016) 675-680.

4.1 INTRODUCTION

Comprehensive two-dimensional gas chromatography (GC × GC) is a powerful technique that can be used to separate complex samples. Comprehensive two-dimensional (2D) separations were conceptually laid out by Giddings [1], and experimentally pioneered in the GC field by Phillips and Liu [2]. In GC × GC, two columns are utilized providing complementary separations which should ultimately produce a separation space with a high degree of informational orthogonality [3]. The two chromatographic columns are coupled using a modulator, which injects portions of eluent containing analyte from the ^{1}D column onto the ^{2}D column with a user defined frequency. While modulator design has become relatively well established, there continues to be a need for the development of high performance modulators that are low cost to purchase and operate, which is the topic of this report.

For the sake of clarity, we define three approaches for GC × GC modulation: thermal modulation, flow modulation, and diaphragm valve-based modulation. Briefly, thermal modulation relies on a low temperature to trap and focus analyte eluting from the ^{1}D column [2,4–9], followed by rapid heating to introduce the analyte in a small volume injection onto the ^{2}D column. Flow modulation closely resembles Dean switching but results in total transfer of analyte

from the ^{1}D column to ^{2}D column [10–12]. In contrast, diaphragm valve-based modulation is distinctively different from the other two forms of modulation since the carrier gas flow rates for the ^{1}D and ^{2}D separations are uncoupled, which provides the advantage to independently optimize the separations on the two dimensions [13–15]. Additional information regarding GC × GC modulation can be found elsewhere [16–21].

Modulation for GC × GC via a diaphragm valve was first reported in 1998, wherein a portion of the ^{1}D column effluent was diverted to the ^{2}D column [13]. Originally, the diaphragm valve had temperature limits of 175 °C due to the temperature sensitive O-rings. This temperature limitation was overcome to some extent by face-mounting the valve on the outside of the oven allowing separations up to 265 °C [14]. Improvements to the diaphragm valve-based modulator were implemented by Seeley and co-workers who used a sample loop to the collect a relatively large fraction of the ^{1}D column effluent which improved detection sensitivity [15]. Recently, diaphragm valve-based GC × GC – FID has been developed with the diaphragm valve O-rings upgraded with a perfluoroelastomer material which allows the valve to be placed directly in the GC oven for separations up to 325 °C [22,23]. This state-of-the-art instrument diaphragm valves are gaining traction as a simple and inexpensive, yet robust, method for modulation in GC × GC. Narrow injection pulses onto the ^{2}D column have been shown to produce high GC × GC peak capacity even with short modulation periods [13,15,24,25]. While the recent work with the high temperature valve-based GC × GC has been with the FID for detection, the next logical step was to investigate coupling of the valve-based GC × GC with time-of-flight mass spectrometry (TOFMS). The use of TOFMS for detection will substantially broaden the scope of using diaphragm valve-based GC × GC.

The TOFMS is ideally suited for GC × GC detection, and in particular diaphragm valve-based GC × GC which produces narrow ^{2}D peaks, because the TOFMS readily collects full spectra at a frequency up to 500 Hz [26,27]. However, when performing diaphragm valve-based GC × GC – TOFMS, special considerations of the flow rate must be taken into account. Commonly for valve-based modulation, the flow rate on the ^{2}D column has been in excess of 10 ml/min which is generally not applicable to GC × GC – TOFMS [14,15,22]. Previously, diaphragm valve-based GC × GC – TOFMS has been performed with a flow rate of 6 ml/min and was limited to separation temperatures of 265 °C using the face mounted valve approach [28]. While this previous study demonstrated that valve-based modulation and TOFMS detection can be used in union, high flow rates on the ^{2}D dimension limit the mainstream applicability to mass spectrometry systems without the use of powerful vacuum pumps. To address this challenge, we targeted use of a maximum flow rate of 3 ml/min for more universal compatibility with TOFMS vacuum pump systems. Herein, we demonstrate the ability to perform diaphragm valve-based GC × GC – TOFMS at a flow rate of 3 ml/min with separation temperatures up to 325 °C with the diaphragm valve mounted directly in the oven. Peak capacity and detection sensitivity are examined. Finally, diesel fuel is evaluated to investigate performance of the instrument on a real-world sample.

4.2 EXPERIMENTAL

The high temperature diaphragm valve was installed on a GC × GC – TOFMS, comprised of an Agilent 6890 N gas chromatograph equipped with an Agilent 7683 autoinjector coupled with a LECO Pegasus III TOFMS (LECO, St. Joseph, MI) as shown in Fig. 4.1. The electron impact ionization voltage for the TOFMS was set to −70 eV, with a detector voltage of 1700 V, and the data was collected from 35 amu to 334 amu at unit mass resolution at a rate of 100 Hz. The ^{1}D column had a RTX-5 ms stationary phase (Restek, Bellefonte, PA, USA), with a 60 m length,

250 µm inner diameter, and 0.5 µm film thickness, and the ^{2}D column had a RTX-200 stationary phase (Restek, Bellefonte, PA, USA), with a 1 m length, 100 µm inner diameter, and 0.1 µm film thickness. A test mixture, which was previously used to evaluate the diaphragm valve [22], was evaluated as listed in Table 4.1 in order to study the instrumental platform. The test mixture contained a wide range of boiling compounds (98−450 °C) with different chemical characteristics (alkanes, alcohols, and polyaromatic hydrocarbons) which allowed for a comprehensive study of the diaphragm valve. A 2 µl volume of the test mixture was injected at a split of 20:1 with an inlet temperature of 325 °C. Ultra-high purity helium (Grade 5, 99.999%) was used as the carrier gas (Praxair, Seattle, WA, USA) at a constant flow of 1.0 ml/min on ^{1}D while the flow on ^{2}D was 3 ml/min. The oven was held at 40 °C for 8 min and ramped at 2 °C/min to 325 °C where it was held for 10 min, with a 5 °C degree offset applied to the ^{2}D oven. The modulation period P_M was 2 s. The high temperature valve was implemented with a 5 µl sample loop and an injection pulse width of 60 ms. Although a larger injection pulse width of ~100 ms is required to totally clear the sample loop (the time in the inject state), a shorter pulse width of 60 ms was found to result in narrower peaks on the ^{2}D dimension with only a modest loss in detection sensitivity.

Diesel fuel, obtained from a local gas station, was evaluated using the high temperature diaphragm valve-based GC × GC – TOFMS. For the diesel separation, a RTX-5 ms stationary phase for the ^{1}D column was also used, but with a 20 m length, 250 µm i.d., and 0.5 µm film thickness. The ^{2}D column had a RTX-200 stationary phase with a 2 m length, 100 µm i.d., and 0.1 µm film thickness. A 2 µl sample of diesel fuel was injected at a split of 20:1 with an inlet temperature of 300 °C. The flow on ^{1}D was 1.0 ml/min while flow on ^{2}D was 3 ml/min. The oven was held at 40 °C for 2 min and ramped at 10 °C/min to 300 °C where it was held for 10 min, and

a 5 °C degree offset was applied to the secondary oven. The modulation period P_M was 1 s, with a 5 µl sample loop and an injection pulse width of 60 ms.

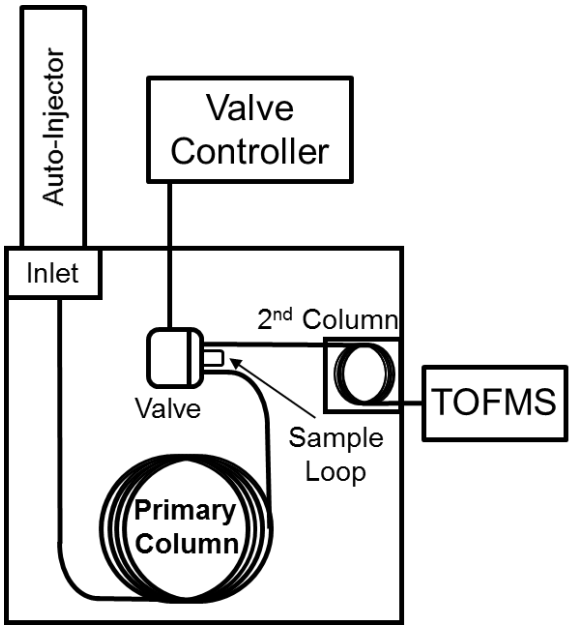


Figure 4.1 High temperature diaphragm valve-based GC × GC–TOFMS instrument schematic. Modulation occurs via a sample loop on the six-port high-speed diaphragm valve.

4.3 RESULTS AND DISCUSSION

In the initial evaluation of the high temperature diaphragm valve-based GC × GC – TOFMS, a chromatogram for the test mixture (Table 4.1) for *m/z* 43 is shown in Fig. 4.2A, demonstrating the ability of the valve to separate a wide range of boiling points (98–450 °C). The earliest eluting analyte (heptane) elutes at a time of 20 min, corresponding to a temperature of 64 °C, while the latest eluting analyte (triacontane), elutes at a time of 133 min, corresponding to a temperature of 306 °C. In order to discuss the peak capacity and sensitivity of the instrument, an enhanced view of docosane is shown in Fig. 4.2B. Docosane, a representative mid-separation analyte has an apparent peak width-at-base ^1W_b of ~10 s (±4 standard deviation peak width). This was determined by fitting a Gaussian function over the data (through the apexes of each of the ^{2}D

peaks in the unfolded vector format) [29–31]. With a modulation period P_M of 2 s, the modulation ratio M_R which is defined as ${}^1W_b/P_M$ was determined to be 5 for docosane [30]. Previous findings using the high temperature diaphragm valve-based GC × GC – FID to determine the precision for retention times and peak widths on the ^{2}D column found that the retention times varied by less than 0.5% and the peak widths varied by less than 3% [22]. Observed results for the high temperature diaphragm valve-based GC × GC – TOFMS were found to be essentially the same and thus are excluded for brevity herein.

Compound	Boiling point (°C)
Alkanes	
Heptane	98
Octane	125
Nonane	151
Decane	174
Undecane	196
Dodecane	216
Tridecane	235
Tetradecane	254
Pentadecane	271
Hexadecane	287
Heptadecane	302
Octadecane	317
Nonadecane	330
Eicosane	343
Heneicosane	357
Docosane	368
Tricosane	380
Tetracosane	391
Pentacosane	401
Hexacosane	412
Heptacosane	422
Octacosane	432
Nonacosane	441
Triacontane	450

Compound	Boiling point (°C)
Alcohols	
Hexanol	157
2-heptanol	159
Octanol	195
Nonanol	215
Decanol	231
Dodecanol	256
Tetradecanol	289
Hexadecanol	344
Octadecanol	210 at 15 mm Hg
Eicosanol	370
Docosanol	180 at 0.22 mm Hg
Tetracosanol	395
Hexacosanol	446
Polyaromatic Hydrocarbons	
Naphthalene	218
Fluorene	295
Phenanthrene	332
Anthracene	340
Fluoranthene	375
Pyrene	404
Chrysene	448

Table 4.1 List of the analytes in the test mixture as well as their boiling points.

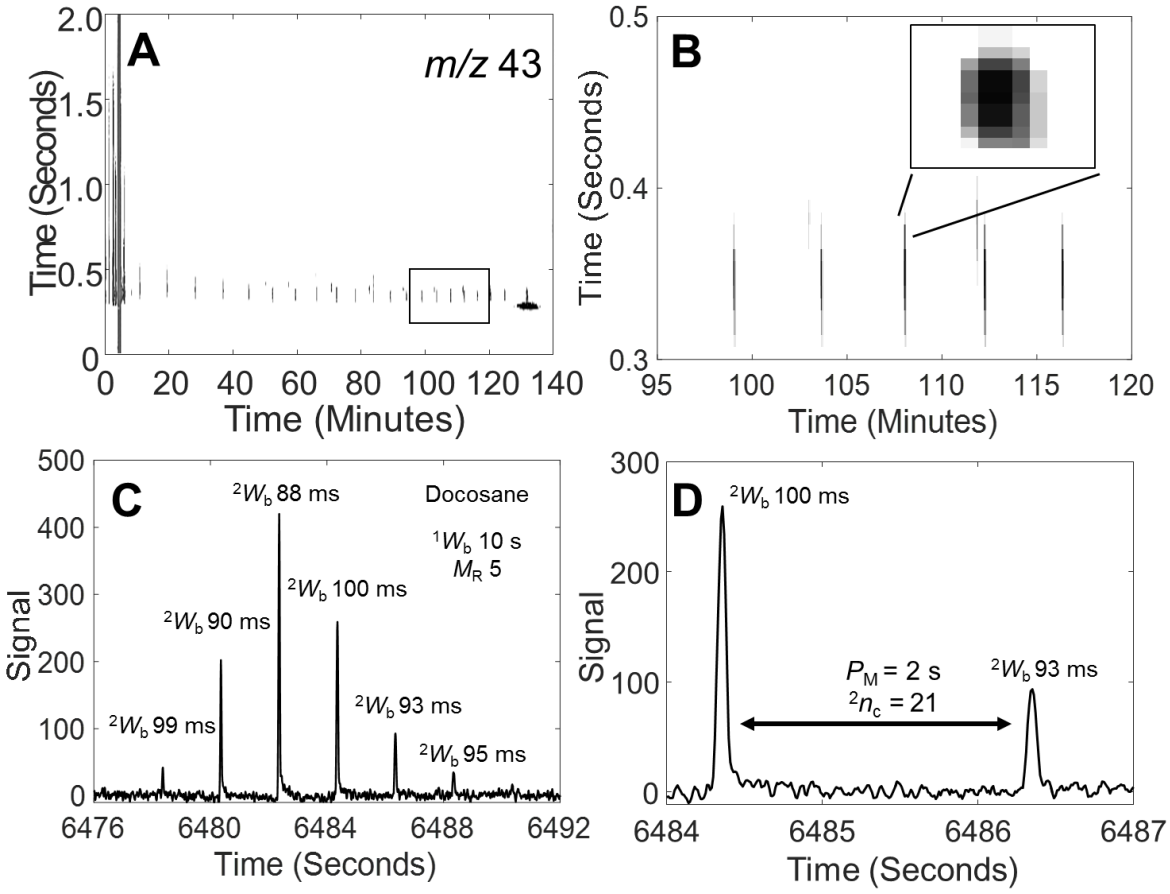


Figure 4.2 (A) Chromatogram of *m/z* 43 is shown. A separation of the test mixture (per Table 4.1) is shown demonstrating the high capability of the valve modulation for a temperature program of 40–325 °C. (B) Enhanced view of a portion of the chromatogram containing eicosane to tetracosane is shown in order to discuss the merits of peak capacity and sensitivity. The 2D view of docosane (C22) is provided as an insert. (C) Unfolded vector form of docosane is shown. Using the 3 ml/min flow rate on ^{2}D, narrow peaks are achieved. (D) Additional zoom in view of docosane is shown. Due to narrow peaks, a large peak capacity is achieved on ^{2}D.

Peak capacity and detection sensitivity discussion is facilitated using the vector form (one dimensional array) of the docosane data provided in Fig. 4.3C and D. Docosane had an average ^2W_b of 94 ms which corresponds to a peak capacity of $^2n_c \sim 20$ on ^{2}D with a P_M of 2 s. The ^2n_c peak capacity was calculated at unit resolution ($R_S = 1$), defined as P_M divided by the ^2W_b. Based on a peak width-at-base on the ^{1}D dimension of 10 s with a run time separation window of 120 min, the ^{1}D peak capacity is $^1n_c \sim 700$. Thus, using docosane as a representative analyte, the ideal 2D peak capacity given by $n_{c,2D} = {}^1n_c \times {}^2n_c$ is hypothetically ~ 14,000. The results in Fig. 4.2 demonstrate that narrow peaks and high peak capacity can be produced on the ^{2}D dimension despite using the 3 ml/min flow rate. In terms of the detection sensitivity for the instrument, the most intense modulation of docosane in Fig. 4.3C, is at a signal-to-noise ratio (S/N) of 100. Docosane was injected at a concentration of 45 ppm, which translates into a limit of quantitation (LOQ) for docosane of 4.5 ppm, and a limit of detection (LOD) of 1.5 ppm. Given the injection split ratio was 20:1, additional improvement in the LOQ and LOD can be readily achieved by reducing the split. It is acknowledged that the detection sensitivity of the valve based instrument is lower than for thermal modulation, but for many applications, especially those in which it is desirable to avoid the expense of cryogens and/or a large split is warranted for the injection onto ^{1}D, then the valve based instrument may very well be the instrumental platform of choice. GC × GC analysis of petroleum based samples (i.e. diesel, RP fuel, or crude oil) often requires high splits for injection onto the ^{1}D column (200:1) in order to avoid ^{2}D column and detector overloading when thermal modulation is used [7,18]. However, when valve-based modulation is implemented with these types of samples, a lower split (20:1) is typically implemented [22–24]. Since roughly 10–40% of the sample is transferred from the ^{1}D column to the ^{2}D column in valve-

based modulation [22] depending upon the separation conditions, roughly the same amount of sample reaches the detector as in thermal modulation.

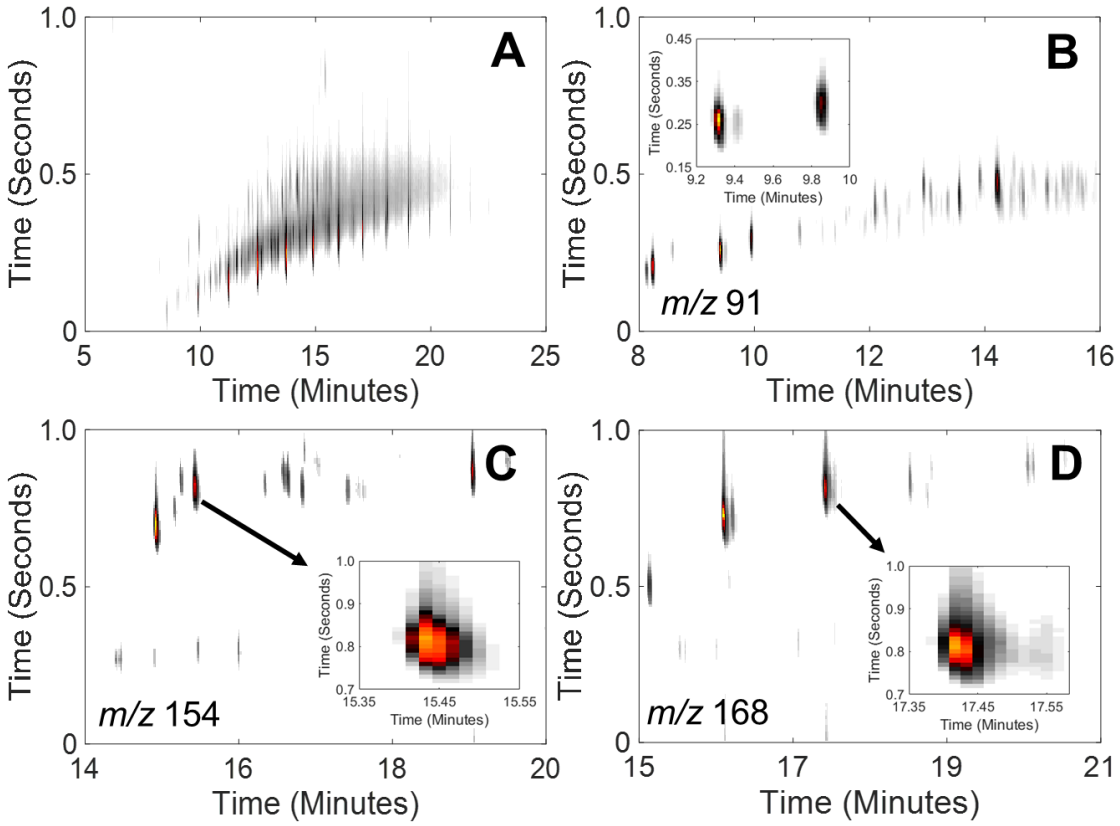


Figure 4.3 (A) TIC chromatogram of diesel fuel is shown. (B) Portion of the chromatogram is shown for *m/z* 91 to highlight the single ring aromatic compounds (C) Portion of the chromatogram is shown for *m/z* 154, to highlight the naphthalene compounds. (D) Portion of the chromatogram is shown for *m/z* 168, which emphasizes additional relatively polar compounds.

The utility of high temperature diaphragm valve-based GC × GC – TOFMS was evaluated using diesel fuel obtained from a local gas station. As shown in the Fig. 4.3A, the total ion chromatogram (TIC) for diesel fuel contains a wide range of compounds ranging over a wide range of volatility. The first analyte elutes on ^{1}D at a ~ 6.5 min, corresponding to a temperature of 85 °C

while the last analyte elutes at ~23.5 min, corresponding to a temperature of 255 °C. A limitation in using the TIC in Fig. 4.3A to visualize the separation is that the diesel fuel components at low concentration are not apparent, since diesel contains compounds that cover an extremely broad concentration range, from a few volume percent to well below the LOD of the instrument, which is about 1 ppm. Therefore, the TIC chromatogram does not clearly show whether or not the instrument is effectively using the 2D peak capacity.

In order to clearly demonstrate use of the 2D peak capacity, Fig. 4.3B–D were prepared, which highlight individual mass channels for selected chemical compound classes. In Fig. 4.3B, using *m/z* 91, a large number of aromatic compounds are indicated. The insert shows two well resolved peaks, propylbenzene (^1t_R = 9.3 min and ^2t_R = 0.25 s) and 1,2,3-trimethylbenzene (^1t_R = 9.8 min and ^2t_R = 0.3 s). Peak widths at base on ^{1}D and ^{2}D were also measured for propylbenzene (^1W_b = 4.6 s, and ^2W_b= 145 ms), and for 1,2,3-trimethylbenzene (^1W_b = 5.0 s, and ^2W_b = 143 ms) for the purpose of calculating the 2D separation space peak capacity. The accompanying mass spectra for propylbenzene (match value = 900) and 1,2,3-trimethylbenzene (match value = 926) are provided in Fig. 4.4A and B, respectively, indicative of the single-ring aromatic compounds being emphasized in Fig. 4.3B. Excellent spectral matches to the NIST MS search program were obtained. The match value is calculated by taking the inner product of the instrumentally obtained mass spectrum relative to the NIST MS data base mass spectrum, with lower *m/z* peaks having less weight than the higher *m/z* peaks. Values are reported from 1 to 1000 where 1000 is a perfect match.

Further evidence for effective use of the 2D peak capacity was obtained. In Fig. 4.3C using *m/z* 156, representative of naphthalene derivatives are highlighted. The insert provides an enhanced view of 2,6-dimethylnaphthalene (^1t_R = 15.44 min and ^2t_R = 0.8 s, and ^1W_b = 6 s, and ^2W_b

= 200 ms), with the mass spectrum provided in Fig. 4.4C (match value = 922), indicative of the naphthalene compounds being highlighted in Fig. 4.3C. Finally, in Fig. 4.3D using *m/z* 168, yet another view of other types of aromatic species is emphasized. The insert in Fig. 4.3D provides an enhanced view of 4-methyl-1,1′-biphenyl (^1t_R = 17.41 min and ^2t_R = 0.8 s, and ^1W_b = 6.3 s, and ^2W_b = 176 ms), with the mass spectrum provided in Fig. 4.4D (match value = 890).

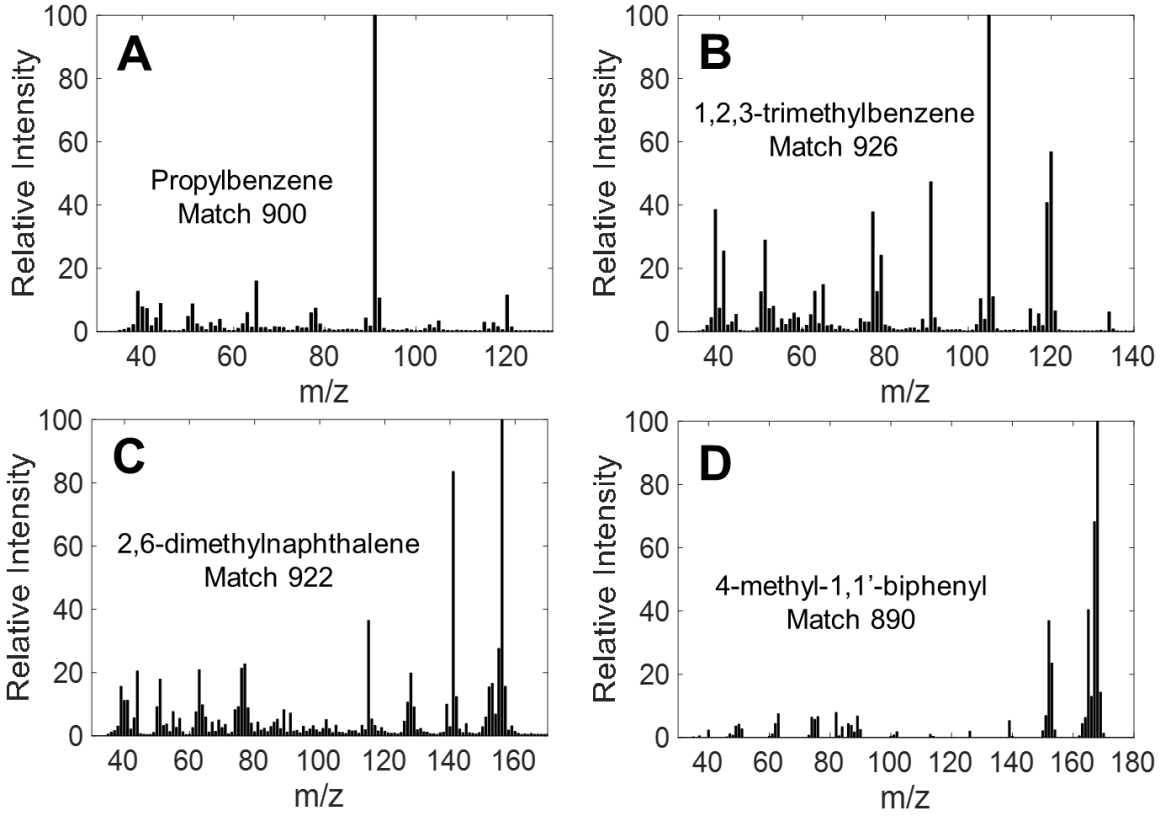


Figure 4.4 (A) Mass spectrum of propylbenzene. (B) Mass spectrum of 1,2,3-trimethylbenzene. (C) Mass spectrum of 2,6-dimethylnaphthalene. (D) Mass spectrum of 4-methyl-1,1′-biphenyl.

The four compounds highlighted in Figs. 4.3B–D and 4.4A–D as well as a couple other analytes (cyclohexane and decane) serve as representative peaks for calculation of the 2D peak capacity. Based upon the ^{1}D and ^{2}D peak width-at-base values summarized in Table 4.2, the average ^{1}D and ^{2}D peak widths were determined to be ^1W_b = 5.4 s, and ^2W_b = 166 ms. Using these

average peak width values, the ^1n_c peak capacity was determined to be 220, based upon a 20 min separation time window (Fig. 4.3A). Likewise, the ^2n_c peak capacity was 6 for the modulation period P_M of 1.0 s. The resulting ideal 2D peak capacity is $n_{c,2D} \sim 1320$. This 2D peak capacity is very consistent, and competitive, with what is generally achieved in the field using thermal modulated GC × GC – TOFMS for ^{1}D run times of only 20 min [30,32].

Compound	^1t_R (min)	^2t_R (s)	^1W_b (s)	^2W_b (ms)
Cyclohexane	5.5	0.89	6.0	170
Propylbenzene	9.3	0.25	4.6	145
1,2,3-trimethylbenzene	9.8	0.30	5.0	143
Decane	10.0	0.12	4.5	160
2,6-dimethylnaphthalene	15.44	0.80	6.0	200
4-methyl-1,1'-biphenyl	17.41	0.80	6.3	176

Table 4.2 Summary of the compounds identified in the diesel fuel for purposes of studying peak capacity in the context of a relatively rapid separation for a complex sample.

4.4 CONCLUSION

A high temperature diaphragm valve has been incorporated as a modulator for GC × GC – TOFMS using a flow rate of 3 ml/min on the ^{2}D dimension. A separation of a test mixture containing alkanes, alcohols, and polyromantic hydrocarbons which encompassed a wide range of boiling points (98–450 °C) was performed. Production of narrow peaks on the ^{2}D dimension (94 ms for docosane) resulted in respectable peak capacity of ~ 20 for a P_M of 2 s. A separation of diesel fuel was performed to demonstrate the applicability of high temperature diaphragm valve-based GC × GC – TOFMS. The instrument performed well for a complex sample such as diesel fuel.

4.5 ACKNOWLEDGEMENTS

We would like to thank Stan Stearns of Valco for providing the high temperature valves.

4.6 REFERENCES

[1] J.C. Giddings, Two-dimensional separations: concept and promise, Anal. Chem., 56 (1984) 1258A–1270A.

[2] Z. Liu, J.B. Phillips, Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, J. Chromatogr. Sci., 29 (1991) 227–231.

[3] Z. Liu, D.G. Patterson Jr., M.L. Lee, Geometric approach to factor analysis for the estimation of orthogonality and practical peak capacity in comprehensive two-dimensional separations, Anal. Chem., 67 (1995) 3840–3845.

[4] J.B. Phillips, R.B. Gaines, J. Blomberg, F.W.M. van der Wielen, J.M. Dimandja, V. Green, J. Granger, D. Patterson, L. Racovalis, H.J. de Geus, J. de Boer, P. Haglund, J. Lipsky, V. Sinha, E.B. Ledford Jr., A robust thermal modulator for comprehensive two-dimensional gas chromatography, J. High Resolut. Chromatogr., 22 (1999) 3–10.

[5] J.B. Phillips, J. Xu, Comprehensive multi-dimensional gas chromatography, J. Chromatogr. A, 703 (1995) 327–334.

[6] P.J. Marriott, R.M. Kinghorn, Longitudinally modulated cryogenic system. A generally applicable approach to solute trapping and mobilization in gas chromatography, Anal. Chem., 69 (1997) 2582–2588.

[7] R.M. Kinghorn, P.J. Marriott, Design and implementation of comprehensive gas chromatography with cryogenic modulation, J. High Resolut. Chromatogr., 23 (2000) 245–252.

[8] E.B. Ledford Jr., C. Billesbach, Jet-cooled thermal modulator for comprehensive multidimensional gas chromatography, J. High Resolut. Chromatogr., 23 (2000) 202–204.

[9] J. Beens, M. Adahchour, R.J.J. Vreuls, K. van Altena, U.A.Th. Brinkman, Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography, J. Chromatogr. A, 919 (2001) 127-132.

[10] J.V. Seeley, N.J. Micyus, S.V. Bandurski, S.K. Seeley, J.D. McCurry, Microfluidic Deans switch for comprehensive two-dimensional gas chromatography, Anal. Chem., 79 (2007) 1840–1847.

[11] B.D. Quimby, J.D. McCurry, W.M. Norman, Capillary flow technology for gas chromatography: reinvigorating a mature analytical discipline, LCGC Peak, 25 (2007) 7–15.

[12] P.A. Bueno, J.V. Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A, 1027 (2004) 3–10.

[13] C.A. Bruckner, B.J. Prazen, R.E. Synovec, Comprehensive two-dimensional high-speed gas chromatography with chemometric analysis, Anal. Chem., 70 (1998) 2796–2804.

[14] A.E. Sinha, K.J. Johnson, B.J. Prazen, S.V. Lucas, C.G. Fraga, R.E. Synovec, Comprehensive two-dimensional gas chromatography of volatile and semi-volatile components using a diaphragm valve-based instrument, J. Chromatogr. A, 983 (2003) 195–204.

[15] J.V. Seeley, F. Kramp, C.J. Hicks, Comprehensive two-dimensional gas chromatography via differential flow modulation, Anal. Chem., 72 (2000) 4346–4352.
[16] P. Marriott, R. Shellie, Principles and applications of comprehensive two-dimensional gas chromatography, Trends Anal. Chem., 21 (2002) 573–583.
[17] J.V. Seeley, Recent advances in flow-controlled multidimensional gas chromatography, J. Chromatogr. A, 1255 (2012) 24–37.
[18] J. Harynuk, T. Gorecki, New liquid nitrogen cryogenic modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A, 1019 (2003) 53–63.
[19] P.J. Marriott, P.D. Morrison, R.A. Shellie, M.S. Dunn, E. Sari, D. Ryan, Recent applications in multidimensional chromatography, LCGC Eur., 16 (2003) 23–31.
[20] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A. Th. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC × GC): I. Introduction and instrumental set-up**,** Trends Anal. Chem., 25 (2006) 438–454.
[21] P.Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, Modulators for comprehensive two-dimensional gas chromatography, Trends Anal. Chem., 30 (2011) 1437–1461.
[22] C.E. Freye, L. Mu, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography, J. Chromatogr. A, 1424 (2015) 127–133.
[23] C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, Partial least squares analysis of rocket Propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection, Talanta, 153 (2016), 203-210
[24] B.J. Prazen, K.J. Johnson, A.W. Weber, R.E. Synovec, Two-dimensional gas chromatography and trilinear partial least squares for the quantitative analysis of aromatic and naphthene content in naphtha, Anal. Chem., 73 (2001) 5677–5682.
[25] J.V. Seeley, F.J. Kramp, K.S. Sharpe, S.K. Seeley, Characterization of gaseous mixtures of organic compounds with dual-secondary column comprehensive two-dimensional gas chromatography, J. Sep. Sci., 25 (2002) 53–59.
[26] J. Dallüge, L.L.P. Van Stee, X. Xu, J. Williams, J. Beens, R.J.J. Vreuls, U.A.Th. Brinkman, Unravelling the composition of very complex samples by comprehensive gas chromatography coupled to time-of-flight mass spectrometry: cigarette smoke, J. Chromatogr. A, 974 (2002) 169-184.
[27] R. Shellie, P.J. Marriott, P. Morrison, Concepts and preliminary observations on the triple-dimensional analysis of complex volatile samples by using GC×GC−TOFMS, Anal. Chem., 73 (2001) 1336-1344.
[28] A.E. Sinha, B. J. Prazen, C.G. Fraga, R.E. Synovec, Valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection: instrumentation and figures-of-merit, J. Chromatogr. A, 1019 (2003) 79-87.
[29] R.E. Murphy, M.R. Schure, J.P. Foley, Effect of sampling rate on resolution in comprehensive two-dimensional liquid chromatography, Anal. Chem., 70 (1998) 1585–1594.
[30] W. Khummueng, J. Harynuk, P.J. Marriott, Modulation ratio in comprehensive two-dimensional gas chromatography, Anal. Chem., 78 (2006) 4578–4587.
[31] W.C. Siegler, B.D. Fitz, J.C. Hoggard, R.E. Synovec, Experimental study of the quantitative precision for valve-based comprehensive two-dimensional gas chromatography, Anal. Chem., 83 (2011) 5190–5196.

[32] W. Khummueng, P.J. Marriott, The nomenclature of comprehensive two-dimensional gas chromatography: defining the modulation ratio (MR), LC–GC Europe, 22 (2009), 38-45.

Chapter 5. Investigation of Ultrafast Separations via a Pulse Flow Valve for MDGC (GC × GC, GC^3)

This chapter was reproduced from C.E. Freye[†], H.D. Bahaghighat[†], R.E. Synovec, "Investigation of Ultrafast Separations via a Pulse Flow Valve for MDGC (GC × GC, GC^3)" *prepared for submission in Journal of Chromatography A*

[†] These authors contributed equally to this work

5.1 INTRODUCTION

Comprehensive two-dimensional gas chromatography (GC × GC) was first conceptually introduced by Giddings [1] and established by Liu and Phillips in 1991 [2]. The transition from the initial instrumentation design to the modern GC × GC platform has yielded a highly efficient and effective instrument for the separation of complex samples [3–9]. In tandem with the development of GC × GC instrumentation, a large amount of effort has been dedicated into developing tools to analyze these information dense datasets [10–15]. There are many reviews written that cover GC × GC instrumental development [16–19] as well as improvements to data analysis techniques [16,17,20,21]. The goal of GC × GC as with any analytical technique is to provide the greatest amount of information within the shortest period of time. For any separation technique, this is normally quantified as the peak capacity or peak capacity production (e.g. peak capacity per unit time). When compared to one-dimensional (1D) GC, GC × GC has roughly a factor of 10 improvement in peak capacity due to addition of the second (^{2}D) dimension [22]. Furthermore, the addition of ^{2}D dimension results in increased chemical selectivity. Analytes that are overlapped on the first (^{1}D) dimension have the opportunity to be resolved on the ^{2}D dimension.

Despite the numerous benefits of GC × GC compared to 1D-GC (increased peak capacity and chemical selectivity), it is still possible to have unresolved analytes. This may be overcome by increasing the overall peak capacity of the instrument, but this may negatively impact the duty cycle (i.e. run time and cool down time) which may be of special concern when multiple samples are analyzed. In order to further enhance the chemical selectivity, it is intriguing to consider higher order instruments especially comprehensive separations. There have been a limited number of comprehensive three-dimensional gas chromatography (GC × GC × GC), herein referred to as GC^3, publications [23–26], but the work serves as an anchor for further refinements to the technique. In the second report [24], the addition of a third dimension (^3D) composed of an ionic liquid stationary phase was leveraged in order to fully resolve several phosphonated compounds (i.e. chemical warfare simulants). More recently, GC^3 was coupled to a TOFMS which allowed for another dimension of selectivity and peak identification [25,26].

While GC^3 is an attractive instrumental platform due to the increased selectivity provided by the three different dimensions, past work has traded peak capacity (or peak capacity production) for an increase in selectivity. In the first report, a peak capacity of 3,500 or peak capacity production of ~45 peaks/min was demonstrated [23]. This was improved by 4-fold in the second report, with approximately the same peak capacity obtained in a 20 minute time period resulting in a peak capacity production of ~180 peaks/min [24]. The third iteration resulted in a similar peak capacity production of ~140 peaks/min and a total peak capacity of 7,000 [25] which is very similar to state-of-the art GC × GC [5,22,27–30]. However, one would expect the peak capacity (and peak capacity production) of GC^3 to exceed that of GC × GC by a significant amount similar to how GC × GC has ~10 times more peak capacity compared to 1D-GC. In order to understand where GC^3 fails to produce the expected increase in peak capacity (and peak capacity production), the

factors that govern 3D peak capacity, $n_{c,3D}$, must be examined. For GC^3, the ideal peak capacity is given by

$$n_{c,3D} = {}^1n_c * {}^2n_c * {}^3n_c \quad (5.1)$$

where ${}^1n_c$, ${}^2n_c$, and ${}^3n_c$ are the ^{1}D, ^{2}D, and ^{3}D peak capacities, respectively. At unit resolution, R_s = 1, equation 1 can be expressed as

$$n_{c,3D} = \frac{{}^1t}{{}^1W} * \frac{{}^2t}{{}^2W}\frac{{}^3t}{{}^3W} \quad (5.2)$$

where ${}^1t$ is the total run time, ${}^2t$ is equivalent to the modulation period, ${}^1P_M$, for the coupling of the ^{1}D to the ^{2}D separations, and ${}^3t$ is equivalent to the modulation period, ${}^2P_M$, for the coupling of the ^{2}D to the ^{3}D separations. The nominal peak width at base (4σ) for each dimension is given by ${}^1W$, ${}^2W$, and ${}^3W$, respectively. In terms of modulation periods, instead of separation run times equation 2 can be expressed as

$$n_{c,3D} = \frac{{}^1t}{{}^1W} * \frac{{}^1P_M}{{}^2W}\frac{{}^2P_M}{{}^3W} \quad (5.3)$$

At the time of the publication of the previous GC^3 work, the fastest modulator published was a diaphragm valve with a P_M of 200 ms which produced peaks ~ 30 ms wide at base [24,31]. In order for GC × GC and GC^3 to be considered comprehensive, the sampling density, ρ_s, defined as the width at base divided by the modulation period, must be greater than 2 [32–34]. Thus peaks sampled by the diaphragm valve must be at a minimum of 400 ms wide. If a peak is 400 ms wide, ${}^2W$, an appropriate modulation period, ${}^1P_M$, might be 3 seconds and thus the peak being sampled must be 6 seconds wide at a minimum. Substituting these numbers in to equation 3, assuming a 30 minute (1800 s) run time

$$n_{c,3D} = \frac{1800\ s}{6\ s} * \frac{3\ s}{400\ ms} * \frac{200\ ms}{30\ ms} = 15{,}000\ or\ 500\ peaks/min \quad (5.4)$$

results in a theoretical maximum peak capacity of 15,000 or 500 peaks/min. However, this equation assumes that peaks on the ^{2}D and ^{3}D separations have constant widths which is not a correct assumption since these separations are considered to be pseudo-isothermal. Thus more appropriate numbers to use would be the average widths for those dimension which might be 600 ms and 40 ms for ^{2}W and ^{3}W, respectively resulting in a peak capacity of 7,500 and peak capacity production of 250 peaks/min. The limiting factor for GC^{3} is the $^{2}P_{M}$ which requires wider than optimum peaks on ^{2}D. Recently, work has been published using a pulse flow valve that can actuate as fast as 50 ms and create peaks as narrow as 15 ms [35]. It is intriguing to use this as a modulator linking the ^{2}D to ^{3}D allowing for optimal separations conditions to be used.

Herein we explore the feasibility of the pulse flow valve to serve as a modulator for GC^{3} and continue to evaluate the pulse flow valve for GC × GC. For the GC × GC instrumentation, we continue to optimize the separations based on the previous work with the pulse flow valve with the overall goal to maximum peak capacity on ^{2}D without reducing the peak capacity of the ^{1}D separation. Additionally, we continue to assess the pulse flow valve as an ultra-fast modulator. The optimized conditions gained from GC × GC study were then applied to GC^{3} instrumentation. We evaluate the GC^{3} instrument using a two test mixtures and a "real world" sample of diesel fuel spiked with 8 compounds.

5.2 EXPERIMENTAL

5.2.1 *Instrumental Summary*

The GC × GC and GC^{3} instruments were both evaluated using a flame ionization detector (FID). The instrumental platforms consisted of an Agilent 6890 GC (Agilent Technologies, Palo Alto, CA, USA). The stock electrometer for the Agilent FID was replaced with a high-speed electrometer built in-house allowing the data to be collected at 100 kHz, with the data boxcar

averaged to 10 kHz. The electrometer was interfaced to a National Instruments data acquisition board, and the resulting data was collected using an in-house written LabVIEW program (National Instruments, Austin, TX, USA). Post-run data processing was performed in MATLAB R2015b (The Mathworks, Inc., Natick, MA, USA). Data was imported into MATLAB and further binned to 1 kHz. Samples were introduced to the instrument via a 7683B auto-injector (Agilent Technologies). Prior to each sample injection, HPLC grade hexane and acetone obtained from Fisher Scientific were used as solvent rinses. Ultra-high purity hydrogen (Grade 5, 99.999%) was used as the carrier gas (Praxair, Seattle, WA, USA). For all instrumental setups an inlet and FID temperature of 250 °C was implemented. All columns were contained within the same oven thus all had the same temperature. Four distinct instrumental setups were implemented (two GC × GC and two GC^3) with each instrumental setup described below.

5.2.2 *GC × GC*

For the two GC × GC evaluations, the GC was fitted with a high-speed pulse valve model 009–1643-900 (Parker Hannifin, Hollis, NH, USA) mounted outside of the oven. The pulsed flow valve was controlled using the same program that was used to collect the FID data. A schematic of the GC × GC instrument is shown in Figure 5.1A. The ^{1}D and ^{2}D columns were linked using a 3-way T-union model MT.5CXS6 (Valco Instruments Company Inc., Houston, TX, USA). An in-house fitting was fabricated to mate the pulse flow valve to 7.24 cm × 1.65 mm copper tubing (Restek, Bellefonte, PA, USA) reduced to 3.81 cm × 0.635 mm steel tubing interfaced with a 2 µl sample loop model CSL2 (Valco Instruments Company Inc.) which was connected to the 3-way T-union.

For the first investigation, the ^{1}D column was a SPB-5 (5% diphenyl/95% dimethyl siloxane) stationary phase: 8 m length × 100 µm inner diameter (I.D.) × 0.1 µm film thickness

(d.f.), and the ^{2}D column was a DB-WAX (polyethylene glycol) stationary phase: 1 m × 100 µm I.D. × 0.1 µm d.f. A test mixture containing a diverse set of 115-compounds was used to evaluate the instrumental platform. The test mixture contained a wide range of boiling points (36–372 °C) with nine chemical compound classes: alkanes, esters, halogenated alkanes, aromatics, alkenes, alkynes, ketones, alcohols, cycloalkanes [25,35]. A 0.5 µl volume of the 115-component test mixture was injected at a split of 100:1. A constant flow rate of 0.8 ml/min was applied. The oven was held at 40 °C for 1 min and ramped at 40 °C/min to 250 °C where it was held for 1 min. A modulation period, P_M, of 500 ms was applied with an injection pulse width of 2 ms (i.e. how long the pulse valve injects carrier gas). The head pressure on the pulse valve was held at 30.5 psig for 1 min and ramped at 1.08 psig/min to 36.2 psig and held for 1 min.

For the second investigation, the ^{1}D column was a SPB-5 stationary phase: 3 m length × 100 µm I.D. × 0.1 µm d.f., and the ^{2}D column was a DB-WAX stationary phase: 1 m × 100 µm I.D. × 0.1 µm d.f. A 0.5 µl volume of a test mixture containing 15 compounds (see Table A.1, Supporting Information) was injected at a split of 100:1. A constant flow rate of 5.3 ml/min was applied. The oven was held at 50 °C for 3 min. The P_M was 75 ms with a 2 ms injection pulse width, and the head pressure on the pulse valve was held at 91 psig.

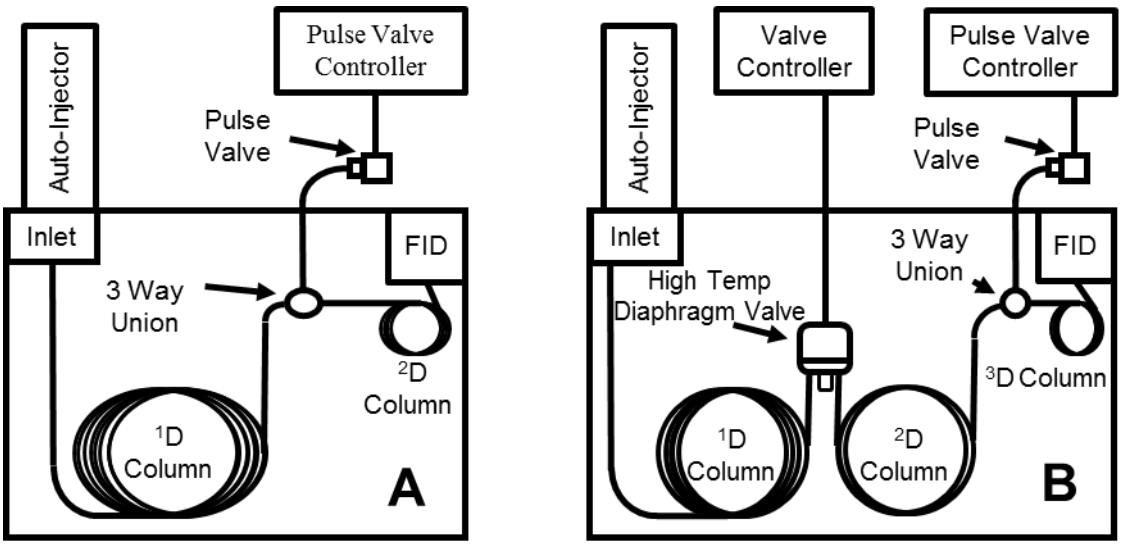


Figure 5.1 (A) Schematic of the major components of the GC × GC - FID instrument. A pulse flow valve was implemented to link the ^{1}D and ^{2}D columns. (B) Schematic of the major components of the GC^{3} - FID instrument. A high-temperature diaphragm valve was utilized as a modulator to link the ^{1}D and ^{2}D columns, and a pulse flow valve was implemented to link the ^{2}D and ^{3}D columns.

5.2.3 *GC^{3}*

For the two GC^{3} instrumental setups, the GC was fitted with a high-speed, six-port, high temperature diaphragm valve VICI model DV-12-1116 T (Valco Instruments Company Inc.) fitted with a 5 µl sample loop which was mounted inside the oven which was used to link ^{1}D and ^{2}D columns [25,27,36,37]. The pulse valve was used to link columns ^{2}D and ^{3}D in an identical manner as described in section 5.2.2. The diaphragm valve, pulse flow valve, and the FID data were all controlled/collected using a LabVIEW program. A schematic of the GC^{3} instrument is shown in Figure 5.1B.

For the third investigation, the ^{1}D column was a RTX-17 Sil MS (50% phenyl/50% methylpolysiloxane) stationary phase: 20 m length × 180 µm I.D. × 0.18 µm d.f., the ^{2}D column was a SPB-5 stationary phase: 6 m length × 100 µm I.D. × 0.1 µm d.f, and the ^{3}D column was a DB-WAX stationary phase: 1 m length × 100 µm I.D. × 0.1 µm d.f. A 0.5 µl volume of an 18 component test mixture (see Table A.2, Supporting Information) was injected at a split of 100:1. A constant flow rate of 0.8 ml/min was applied to ^{1}D. The ^{2}D and ^{3}D columns were controlled via the auxiliary pressure controller with a pressure of 100 psig which corresponds to a flow rate of ~6.3 ml/min. The oven was held at 50 °C for 3 min. A $^{1}P_{M}$ of 1 sec with an injection pulse width of 20 ms applied to the diaphragm valve and a $^{2}P_{M}$ of 50 ms with an injection pulse width of 2 ms was applied to the pulse valve. The head pressure on the pulse valve was held at 83 psig.

For the fourth investigation, the ^{1}D column was a RTX-17 Sil MS stationary phase: 20 m length × 180 µm I.D. × 0.18 µm d.f., the ^{2}D column was a SPB-5 stationary phase: 2 m length × 100 µm I.D. × 0.1 µm d.f, and the ^{3}D column was a DB-WAX stationary phase: 0.75 m length × 100 µm I.D. × 0.1 µm d.f. 1 µl of the 115-component mixture was injected at a split of 60:1. A constant flow rate of 0.7 ml/min was applied to ^{1}D. The ^{2}D and ^{3}D columns were controlled via the auxiliary pressure controller under a ramped pressure program held at 39.2 psig for 1.5 min, ramped at 2.55 psig/min to 65.9 psig, and held at the final pressure for 1 minute. This pressure ramp corresponds to a flow rate of ~3 ml/min. The oven was held at 40 °C for 1.5 min, ramped at 20 °C/min to 250 °C, and held for 1 min. A $^{1}P_{M}$ of 1.2 sec with an injection pulse width of 20 ms applied to the diaphragm valve and a $^{2}P_{M}$ of 60 ms with an injection pulse width of 2 ms was applied to the pulse valve. The head pressure on the pulse valve was held at 36.9 psig for 1.5 min, ramped at 2.55 psig/min to 63.7 psig and held for 1 min. Diesel fuel spiked with 8 compounds (see

Table A.3, Supporting Information) was also evaluated using the same conditions as noted above except the split was reduced to 20:1.

5.3 RESULTS AND DISCUSSION

5.3.1 *GC × GC*

The GC × GC - FID chromatogram for the 115-component test mixture using a modulation period, P_M, of 500 ms in the raw vector form is shown in Fig 5.2A. Three representative analytes (Methyl Decanoate, Pentadecane, Dodecanol) are shown in Fig. 5.2B which range from relatively unretained on ^{2}D to moderately retained. In the raw vector format, the sharp decreases in signal is representative of the ^{2}D analyte response of interest, nominally in the form of an "error" function, followed by a slow decay of the analyte response back to the original ^{1}D concentration profile. This ^{2}D analyte response is created by rapid actuation of the pulse flow valve in combination with the three-way union and coupled with the tubing assembly (see Fig 5.1A). Reducing the tubing assembly from a large inner diameter to a smaller inner diameter results in a pressure wave in the form of an exponentially modified Gaussian with a very sharp leading edge. Using the same method as previously reported [35], a 3-step method involving differentiation, inversion, and baseline correction was implemented to convert the data from the raw form into the final processed form which appears analogous to a "traditional" vector GC × GC chromatogram. Figure A.1 (Supporting Information) demonstrates the data processing method. Figure 5.2C shows the processed vector chromatogram of the data shown in Figure 5.2B.

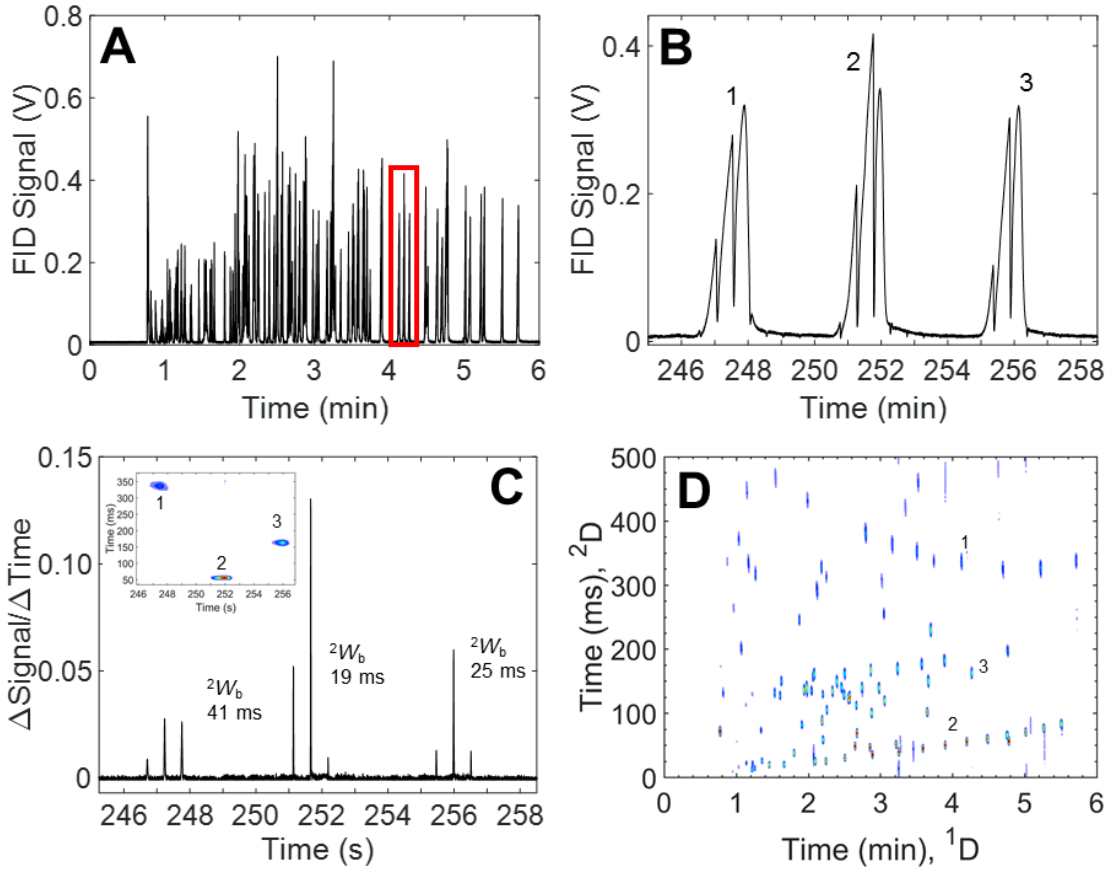


Figure 5.2 (A) Raw (preprocessed) chromatogram of the 115 component mixture collected using a modulation period P_M of 500 ms. The highlighted subset shown contains three non-overlapped analytes (1: Methyl Decanoate, 2: Pentadecane, 3: Dodecanol). (B) Enhanced view of the three non-overlapped analytes is shown. (C) Processed vector chromatogram of the three analytes with width of base annotated. The insert shows the 2D chromatogram of the three analytes. (D) GC × GC - FID chromatogram of the 115 component mixture. Using the three step data processing method, the raw chromatogram shown in (A) has been converted into a "traditional" 2D plot.

The ^2W_b is annotated for the three analytes. Since a polar column (DB-WAX) was implemented as the ^{2}D dimension, the most retained analyte should be methyl decanoate, an ester, which has an apparent ^2W_b of 41 ms. The next most retained analyte is dodecanol with an apparent ^2W_b of 25 ms, and the least retained analyte is pentadecane which has an apparent ^2W_b of 19 ms. The ^2W_b of each analyte is dependent on the slope of the "error" function, and the ^{2}D retention time, ^2t_R, is defined by the inflection point of the "error" function. The insert in Figure 5.2C shows the GC × GC chromatogram of these three analytes which has been registered by 200 ms on the ^{2}D to remove the dead time. Figure 5.2D shows the GC × GC chromatogram of the data processed in Figure 5.2A. The data was registered on ^{2}D by 200 ms to remove the dead time.

In order to accurately measure the ^{2}D peak capacity (^2n_c), 12 representative analytes that spanned the entire ^{2}D separation were selected (see Figure A.2, Supporting Information). The ^2t_R and the ^2W_b for each analyte was measured and an equation was fitted to those values [34–36]. Compared to the previous report [35], the slope of the equation is ~40% less which is due to the use of micro-bore columns and optimization of the pulse flow valve. Using the following recursive equation, the ^2n_c was calculated.

$$R_s = \frac{t_{R.n} - t_{R,n-1}}{0.5(W_{b,n} - W_{b,n-1})} = 1 \qquad (5.5)$$

The equation obtained from Figure A.2 (Supporting Information) was subsituted into equation 5.5 for both $W_{b,n}$ and $W_{b,n-1}$ and rearranging allows the calculation of the retention times of successive peaks (and R_s = 1) based on the measured dead time, $t_{R,(n=0)}$ on ^{2}D [35,38,39]. These retention times were used to calculate apparent ^2W_b from the experimentally observed linear relationship between ^2t_R and apparent ^2W_b. Individual Gaussian peaks of equivalent area were modeled based on these peak widths and retention times, then concatenated to give a continuous chromatogram (see Figure A.2, Supporting Information).

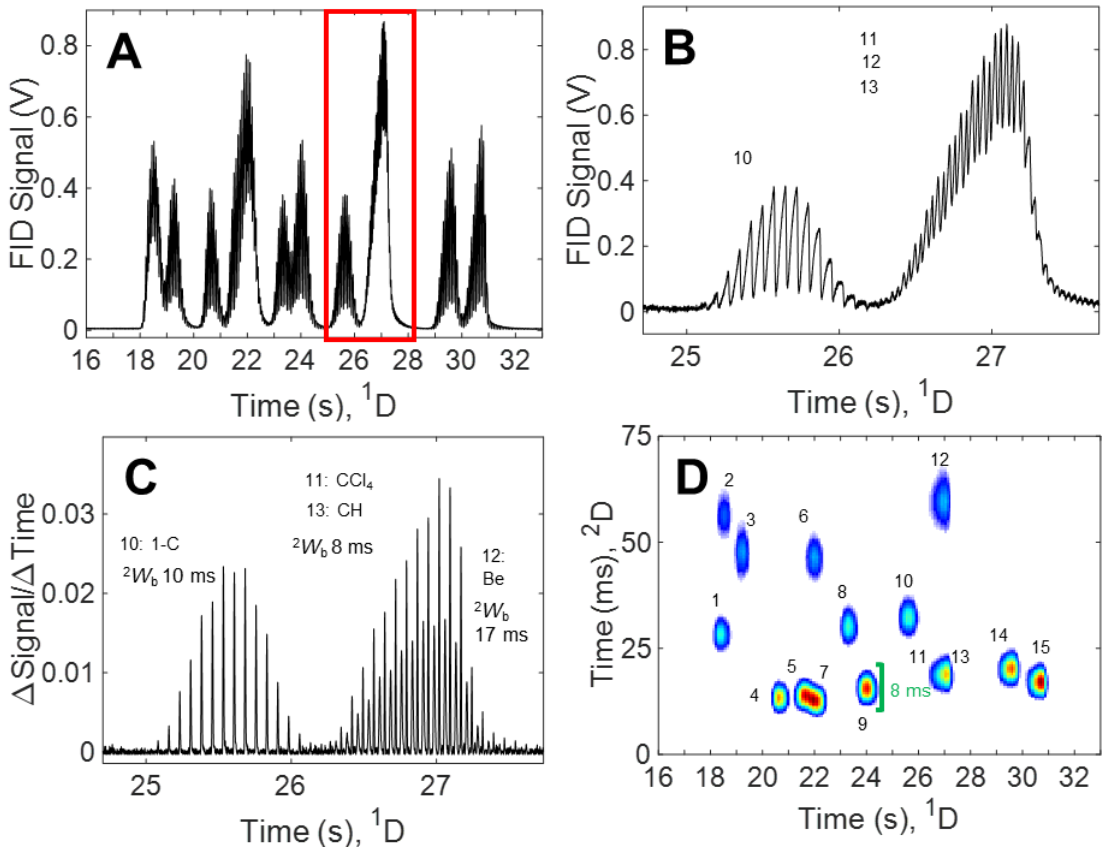


Figure 5.3 (A) Raw (preprocessed) chromatogram of a GC × GC - FID separation of a 15-component low boiling point mixture collected using a modulation period P_M of 75 ms. The highlighted subset is shown in figure (B). (B) Close up view of raw chromatogram containing four analytes numbered in elution order (10: 1-Chlorobutane, 11: Carbon Tetrachloride, 12: Benzene, 13: Cyclohexane). Analytes 11 and 13 are observed to be overlapped in both dimensions which can be seen in (D) (C) Processed vector chromatogram of the four analytes with width of base annotated. (D) GC × GC - FID chromatogram of the 15 component mixture. Using the three step Data processing method, the raw chromatogram shown in (A) has been converted into a "traditional" 2D plot. The names of the compounds can be found in Table A.1. The ^{2}D peaks are extremely narrow, with an average ^2W_b of 10 ms. However, the ^{1}D peaks have been oversampled with a ρ_S of ~12 as seen in (C).

The resulting $^{2}n_c$ was calculated to be 20. Using the same 12 analytes, the average $^{1}W_b$ was calculated to be ~1 s resulting in a ^{1}D peak capacity ($^{1}n_c$) of 360 (t_{sep} = 6 min). Thus the 2D peak capacity, $n_{c,2D,}$ is ~7200 or a peak capacity production of ~1200 peaks/min for this separation.

While pulse flow valve using a P_M of 500 results in extremely high peak capacity production, investigation of the potential as a high speed modulator is warranted. Using a P_M of 75 ms, a 15 component mixture (see Table A.1, Supporting Information) was evaluated. Figure 5.3A shows the raw vector chromatogram of the separation with a t_{sep} of ~13 s. Figure 5.3B shows an enhanced view of Figure 5.3A where 4 analytes are present. One analyte (10: 1-chlorobutane) is totally resolved on ^{1}D while three analytes (11: carbon tetrachloride, 12: benzene, and 13: cyclohexane) are severely overlapped. The processed vector chromatogram is shown in Figure 5.3C with the $^{2}W_b$ annotated. Figure 5.4D shows the 2D chromatogram obtained from processing the data in Figure 5.4A. The narrowest peaks measured (analytes 4, 5, 7, and 9) had $^{2}W_b$ of 8 ms. The $^{2}n_c$ was calculated to be 7.5 and the $^{1}n_c$ was calculated to be ~14.5 ($^{1}W_{b,avg}$ ~ 0.9 s and t_{sep} ~ 13 s) resulting in a $n_{c,2D}$ of ~110 or a peak capacity production of ~510 peaks/min. However, the ^{1}D peaks were significantly oversampled with a sampling density, ρ_s, of ~12. In order to reduce the ρ_s to an appropriate number (~2-4), one could use a modified injection source [38,40–43] or perhaps more interesting would be to sample analytes with pulse flow valve as the second modulator in a GC^3 instrument.

5.3.2 *GC^3*

Building from past research of high temperature diaphragm valve-based GC × GC [27,36] and GC^3 [25,26] it was considered ideal to combine the known attributes of the high temperature diaphragm valve modulator with pulse flow valve modulator. Figure 5.1B shows the instrumental schematic for the GC^3 separations.

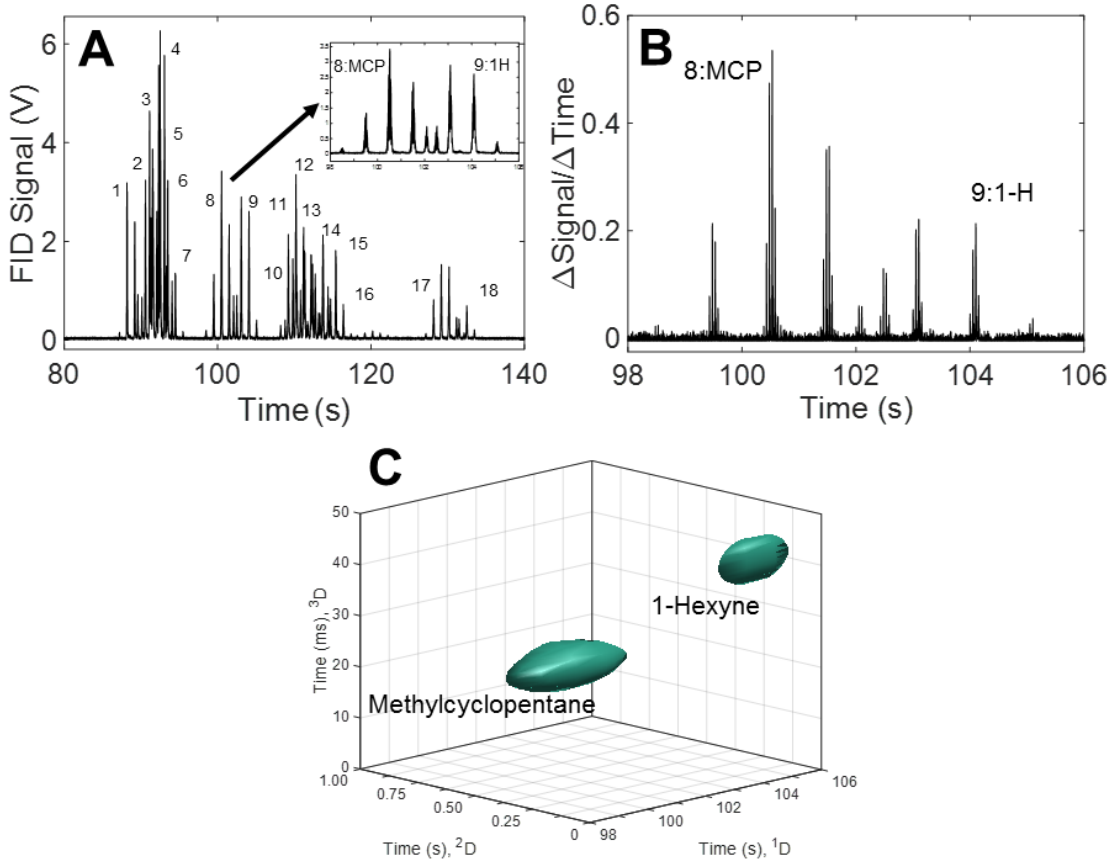


Figure 5.4 (A) The raw (pre-processed) vector chromatogram of a GC^3 - FID separation of the 18 component mixture. The ${}^1P_M$ was 1 sec and the ${}^2P_M$ was 50 ms. An insert of two analytes (8: methylcyclopentane and 9: 1-Hexyne) are shown which are slightly overlapped on the ${}^1D$ dimension. (B) An enhanced view of the processed vector chromatogram is shown for the two analytes highlighted in (A). (C) Isosurface plot of the region depicted in (B).

Figure 5.4A shows the raw vector chromatogram of an 18 component mixture (see Table A.2, Supporting Information) that was collected using a ${}^1P_M$ of 1.0 seconds and a ${}^2P_M$ of 50 ms. The elution order was determined by a series of individual injections of the analytes under the same

separation conditions of the 18 analyte mixture. A processed vector chromatogram of the 18 individual analytes that are overlaid onto one chromatogram is shown in Figure A.3 (Supporting Information). Figure 5.4B shows the processed vector chromatogram of two analytes (8: methylcyclopentane and 9: 1-hexyne) that exhibit a slight overlap on the ^{1}D dimension. In order to demonstrate the selectivity provided by the addition of two dimensions (^{2}D and ^{3}D), Figure 5.4C shows an isosurface plot of the same region shown in Figure 5.4B. The data was registered by 10 ms on the ^{3}D to remove the dead time.

To further explore the selectivity provided by the three separation dimensions, the data shown in Figure 5.4A has been plotted in Figure 5.5A and 5.5B which show the ^{1}D vs ^{2}D where the data has been summed along the ^{3}D and ^{1}D vs ^{3}D where the data has been summed along the ^{2}D, respectively. The chemical selectivity provided by the 3 dimensions can easily be seen when elution pattern of Figure 5.5A is compared to Figure 5.5B. Analytes that are not resolved or only partially resolved on ^{1}D are either separated on the ^{2}D or ^{3}D dimensions. To further explore the selectivity of the GC^{3} system, the ^{2}D vs ^{3}D chromatogram has been prepared in Figure 5.5C where the data was summed along the ^{1}D. Figure 5.5D is a 3-D isosurface plot of the entire separation. In order to calculate the peak capacity of the GC^{3} instrument, three representative analytes were selected (n-hexane, 1-hexyne, methylcyclopentane). These 3 analytes span a wide range of retention times on all three dimensions and thus deemed appropriate in calculate the peak capacity. The figures of merit for the 3 analytes are shown in Table 5.1. For the three representative analytes, the total peak capacity (1-minute total time), $n_{c,3D}$, ranged from 772 to 914 with an average of ~825 or a peak capacity production of 825 peaks/min, which is a 4.5 times improvement over previously published GC^{3} values [24,25]. These results and were based on a simple mixture separated

isothermally which in itself is impressive; however, further work on a more complex mixture that required temperature ramping is needed.

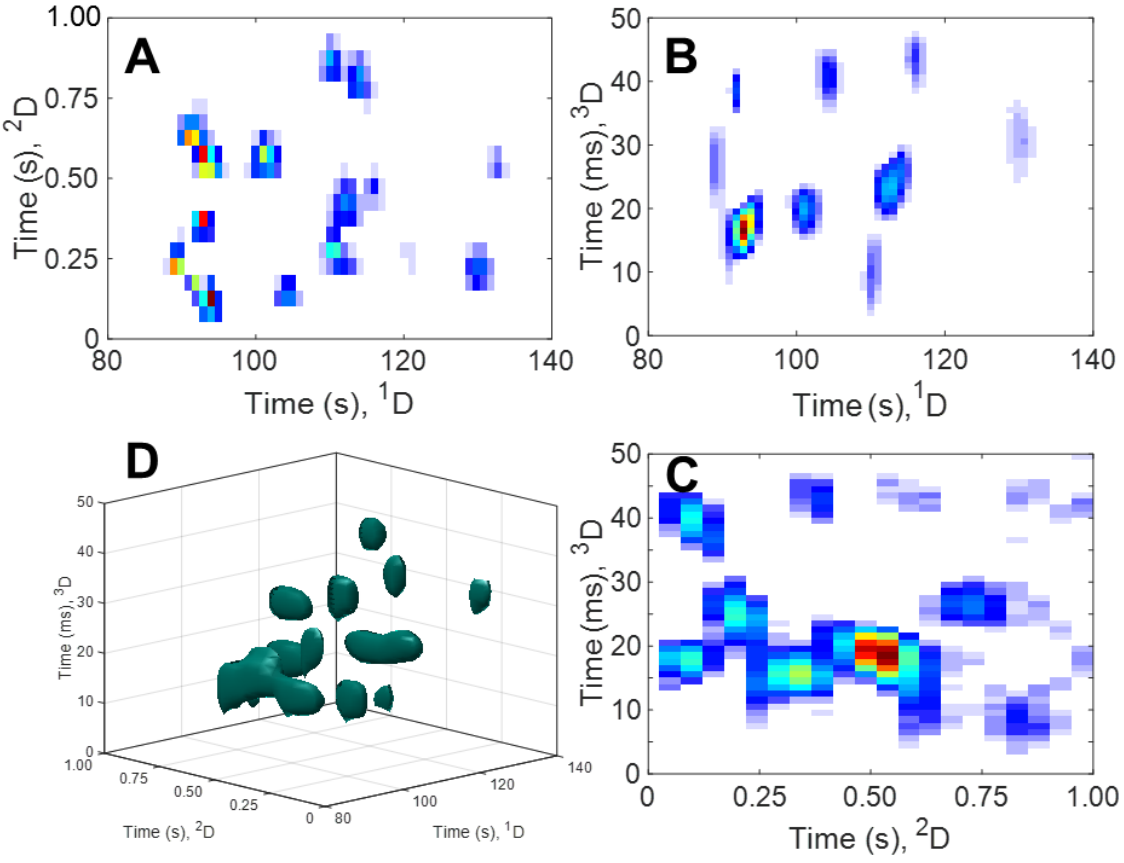


Figure 5.5 In order to show the benefits of the chemical selectivity of the GC^3 - FID system, three contour plots and an isosurface plot have been prepared for the separation of the 18 component mixture. (A) Contour plot of ^2D vs ^1D summed along the ^3D. (B) Contour plot of ^3D vs ^1D summed along the ^2D. (C) Contour plot of ^3D vs ^2D summed along the ^1D. (D) Isosurface plot. An example of the chemical selectivity provided by the three different columns can be seen in A and B. Analytes (8) methylcyclopentane and (9) 1-hexyne have a dramatic shift in retention of the non-polar ^2D and polar ^3D columns. This shift in affinity for ^2D and ^3D columns is highlighted in Figure 5.3.

	n-Hexane	1-Hexyne	Methylcyclopentane
${}^1t_R$ (min)	1.54	1.68	1.64
${}^1W_b$ (sec)	2.26	2.28	2.41
${}^1\rho_s$	2.26	2.28	2.41
${}^1n_c$	26.50	26.36	24.89
${}^2t_R$ (sec)	0.19	0.47	0.35
${}^2W_b$ (msec)	177	128	171
${}^2\rho_s$	3.54	2.56	3.43
${}^2n_c$	5.65	7.83	5.84
${}^3t_R$ (msec)	27.5	3.2	31.8
${}^3W_b$ (msec)	8.2	13.1	9.4
${}^3n_c$	6.10	3.83	5.31
$n_{c,3D}$	914	790	772
Peaks/min	914	790	772

Table 5.1 Chromatographic peak measurements and figures of merit for three representative compounds in the 18 component mix.

In order to continue to evaluate the merits of the GC^3 system, the same 115 component mixture was assessed using optimized experimental conditions (see Section 5.3). Figure 5.6A shows the processed vector chromatogram that was collected using a ${}^1P_M$ of 1.2 seconds and a ${}^2P_M$ of 60 ms. In order to see the selectivity provided by three dimensions, Figure 5.6B and 5.6C have been created which show the ${}^1D$ vs ${}^3D$ where the data has been summed along the ${}^2D$ and ${}^1D$ vs ${}^2D$ where the data has been summed along the ${}^3D$, respectively. The chemical selectivity provided by the 3 dimensions can be especially be seen at ~ 4 to 6 minute range on ${}^1D$. Analytes that are not resolved or only partially resolved on ${}^1D$ are either separated on the ${}^2D$ or ${}^3D$ dimensions. To further explore the selectivity of the GC^3 system, the ${}^2D$ vs ${}^3D$ chromatogram has been prepared (see Fig A.5A, Supporting Information). When the all the data along the ${}^1D$ dimension is summed, the chemical selectivity that is provided by this view disappears. This is because the analyst is

"looking down the tube" and thus it is difficult to extract useful information. Thus 3 ^{2}D vs ^{3}D chromatograms have also been prepared (see Fig A.5B-D, Supporting Information) that span a range of 60 seconds (same as GC^{3} separation of the 18 component mixture). Figure 5.6D shows the isosurface plot of the 115 component mixture. Four analytes listed in Table 5.2 (see Fig A.6, Supporting Information) have been selected to demonstrate the figures of merit for the GC^{3} system. While this is a small portion of the 115 component mixture, these 4 analytes span a wide range of times on all three dimensions and thus deemed appropriate to quote the figures of merit. The method for determining these values was the same as those used in the evaluation of the 18 component mixture. For the 4 representative analytes, the total peak capacity, $n_{c,3D}$, ranged from 5600 to 19,000 with an average of ~10,700. Perhaps more importantly is the peak capacity production which ranges from 520 to 1,800 with an average of ~1000 peaks/min which is a significant improvement over previously published GC^{3} values of 140 and 180 peaks/min [24,25].

	Analyte 1	**Analyte 2**	**Analyte 3**	**Analyte 4**
$^{1}t_{R}$ (min)	1.19	2.03	5.46	6.01
$^{1}W_{b}$ (sec)	2.45	2.47	2.86	4.98
$^{1}\rho_{s}$	2.04	2.06	2.38	4.15
$^{1}n_{c}$	267	263	229	130
$^{2}t_{R}$ (sec)	0.42	0.28	0.29	1.02
$^{2}W_{b}$ (msec)	131	136	125	122
$^{2}\rho_{s}$	2.18	2.26	2.09	2.03
$^{2}n_{c}$	9.7	8.8	9.7	9.9
$^{3}t_{R}$ (msec)	22.5	37.2	17.7	13.0
$^{3}W_{b}$ (msec)	7.9	17.2	14.2	13.8
$^{3}n_{c}$	7.58	3.50	4.23	4.35
$n_{c,3D}$	19622	8125	9328	5597
Peaks/min	1811	750	861	517

Table 5.2 Chromatographic peak measurements and figures of merit for the 115 component test mixture. See supporting information Fig A.6 for location of the analytes in the chromatograms.

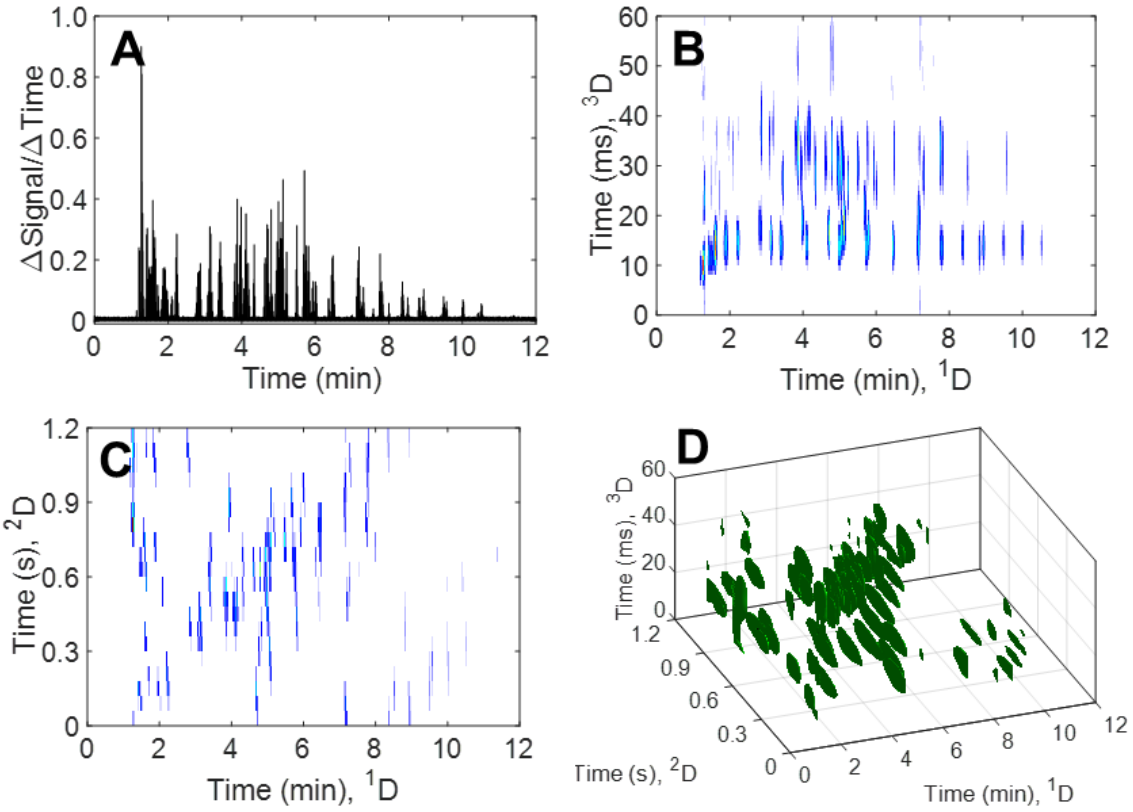


Figure 5.6 (A) Vector chromatogram (processed data) of the 115 component test mixture collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms. (B) Plot of ^{1}D vs ^{3}D summed along ^{2}D. (C) Plot of ^{1}D vs ^{2}D summed along ^{3}D. (D) Isosurface plot of the 115 component mix. For ^{2}D vs ^{3}D see supporting information Fig A.5.

To further show the selectivity provided by the GC^{3} system, diesel fuel was spiked with 8 non-native analytes. Figure 5.7A shows the processed vector chromatogram of the diesel fuel. In this view, it is possible to see the characteristic "petroleum hump" which can be totally resolved using GC^{3}. For this analysis, the conditions were optimized to provide the most separation for the 8 spike analytes instead of attempting to totally resolve the matrix (diesel) background. An isosurface plot shown in Figure 5.7B, has been prepared to shows the selectivity for the 8 spiked analytes. In order visual the data, analytes that have a ^{3}D retention times from 0 to 25 ms are colored blue, analytes

that have a ^{3}D retention times from 25 to 40 ms are colored green, and analytes that have a ^{3}D retention times from 40 to 60 ms are colored red. The blue color roughly corresponds to the location of the alkanes and cycloalkanes and the green color roughly corresponds to the location of the aromatics present in diesel fuel. The 8 spiked analytes were highly retained on ^{3}D and have ^{3}D retention times >40 ms and thus are colored red. In order to visualize the diesel fuel in a planar fashion analogous to a GC × GC chromatogram, Figure A.7 (Supporting Information) has been prepared.

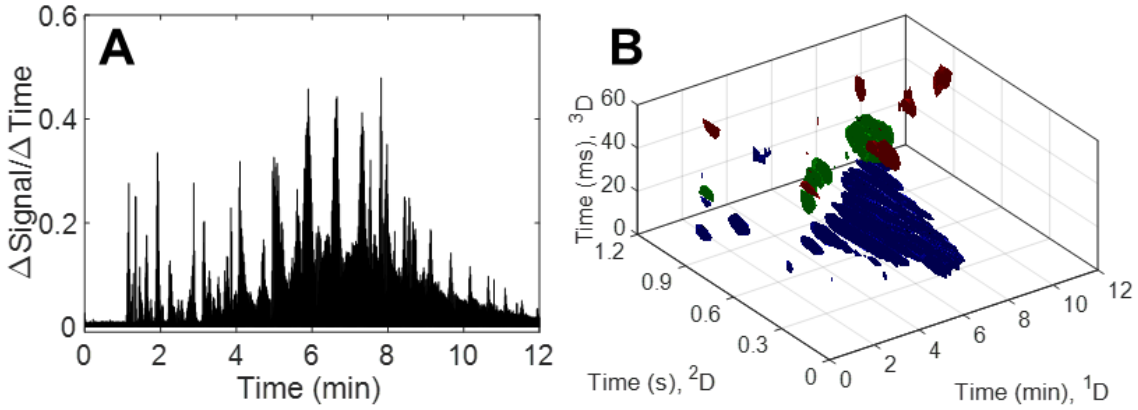


Figure 5.7 (A) Vector chromatogram (processed data) of the diesel that has been spiked with 8 components collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms. (B) Isosurface plot of diesel spiked with the 8 components. Analytes that have a ^{3}D retention time from 0 to 25 ms are colored blue, analytes that have ^{3}D retention time from 25 to 40 ms are colored green, and analytes that have ^{3}D retention times from 40 to 60 ms are colored red. All 8 spiked compounds have ^{3}D retention times greater than 40 ms and thus are colored red. For ^{1}D vs ^{2}D, ^{1}D vs ^{3}D, and ^{2}D vs ^{3}D chromatograms, see supporting information Fig A.7.

5.4 CONCLUSION

Past applications of GC^3 instrumentation demonstrated great promise based on the increased selectivity provided for separations that was superior to GC × GC, but this increase in selectivity came at the cost of peak capacity and peak capacity production. It has been demonstrated that the use of a high temperature diaphragm valve in conjunction with a recently introduced pulse flow valve, it is now possible to take advantage of the increased selectivity of a GC^3 instrument, while achieving a peak capacity production of ~1000 peaks/min. Future studies will be focused on applying these instrumental improvements with spectral data, to create an ultra-fast 4D data structure to address complex separation mixtures.

5.5 ACKNOWLEDGEMENTS

We would like to thank Stan Stearns of Valco for providing the high temperature valves.

5.6 SUPPORTING INFORMATION

Figures A.1 to A.7 and Table A.1 to A.3 can be found in Appendix A.

5.7 REFERENCES

[1] J.C. Giddings, Unified separation science, 1st ed., Wiley-Interscience, 1991.

[2] Z. Liu, J. B. Phillips, Comprehensive Two-Dimensional Gas Chromatography using an On-Column Thermal Modulator Interface, Journal of Chromatographic Science. 29 (1991) 227–231. doi:10.1093/chromsci/29.6.227.

[3] V. Abrahamsson, N. Ristic, K. Franz, K. Van Geem, Comprehensive two-dimensional gas chromatography in combination with pixel-based analysis for fouling tendency prediction, Journal of Chromatography A. 1501 (2017) 89–98. doi:10.1016/j.chroma.2017.04.021.

[4] H.D. Bean, J.-M.D. Dimandja, J.E. Hill, Bacterial volatile discovery using solid phase microextraction and comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, Journal of Chromatography B. 901 (2012) 41–46. doi:10.1016/j.jchromb.2012.05.038.

[5] L.M. Dubois, K.A. Perrault, P.-H. Stefanuto, S. Koschinski, M. Edwards, L. McGregor, J.-F. Focant, Thermal desorption comprehensive two-dimensional gas chromatography coupled to variable-energy electron ionization time-of-flight mass spectrometry for monitoring subtle changes in volatile organic compound profiles of human blood, Journal of Chromatography A. 1501 (2017) 117–127. doi:10.1016/j.chroma.2017.04.026.

[6] A. Giri, M. Coutriade, A. Racaud, K. Okuda, J. Dane, R.B. Cody, J.-F. Focant, Molecular Characterization of Volatiles and Petrochemical Base Oils by Photo-Ionization GC×GC-TOF-MS, Analytical Chemistry. 89 (2017) 5395–5403. doi:10.1021/acs.analchem.7b00124.

[7] H. Potgieter, R. Bekker, J. Beigley, E. Rohwer, Analysis of oxidised heavy paraffininc products by high temperature comprehensive two-dimensional gas chromatography, Journal of Chromatography A. 1509 (2017) 123–131. doi:10.1016/j.chroma.2017.06.046.

[8] S. Prebihalo, A. Brockman, J. Cochran, F.L. Dorman, Determination of emerging contaminants in wastewater utilizing comprehensive two-dimensional gas-chromatography coupled with time-of-flight mass spectrometry, Journal of Chromatography A. 1419 (2015) 109–115. doi:10.1016/j.chroma.2015.09.080.

[9] P.-H. Stefanuto, K.A. Perrault, L.M. Dubois, B. L'Homme, C. Allen, C. Loughnane, N. Ochiai, J.-F. Focant, Advanced method optimization for volatile aroma profiling of beer using two-dimensional gas chromatography time-of-flight mass spectrometry, Journal of Chromatography A. 1507 (2017) 45–52. doi:10.1016/j.chroma.2017.05.064.

[10] S.L. Forbes, A.N. Troobnikoff, M. Ueland, K.D. Nizio, K.A. Perrault, Profiling the decomposition odour at the grave surface before and after probing, Forensic Science International. 259 (2016) 193–199. doi:10.1016/j.forsciint.2015.12.038.

[11] B. Kehimkar, J.C. Hoggard, L.C. Marney, M.C. Billingsley, C.G. Fraga, T.J. Bruno, R.E. Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, Journal of Chromatography A. 1327 (2014) 132–140. doi:10.1016/j.chroma.2013.12.060.

[12] B.A. Weggler, T. Gröger, R. Zimmermann, Advanced scripting for the automated profiling of two-dimensional gas chromatography-time-of-flight mass spectrometry data from combustion aerosol, Journal of Chromatography A. 1364 (2014) 241–248. doi:10.1016/j.chroma.2014.08.091.

[13] J. Omar, M. Olivares, J.M. Amigo, N. Etxebarria, Resolution of co-eluting compounds of Cannabis Sativa in comprehensive two-dimensional gas chromatography/mass spectrometry detection with Multivariate Curve Resolution-Alternating Least Squares, Talanta. 121 (2014) 273–280. doi:10.1016/j.talanta.2013.12.044.

[14] C. Hurtado, H. Parastar, V. Matamoros, B. Piña, R. Tauler, J.M. Bayona, Linking the morphological and metabolomic response of Lactuca sativa L exposed to emerging contaminants using GC × GC-MS and chemometric tools, Scientific Reports. 7 (2017). doi:10.1038/s41598-017-06773-0.

[15] C.E. Freye, N.R. Moore, R.E. Synovec, Enhancing the chemical selectivity in discovery-based analysis with tandem ionization time-of-flight mass spectrometry detection for comprehensive two-dimensional gas chromatography, Journal of Chromatography A. 1537 (2018) 99–108. doi:10.1016/j.chroma.2018.01.008.

[16] S.E. Prebihalo, K.L. Berrier, C.E. Freye, H.D. Bahaghighat, N.R. Moore, D.K. Pinkerton, R.E. Synovec, Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications, Anal. Chem. 90 (2018) 505–532. doi:10.1021/acs.analchem.7b04226.

[17] J.V. Seeley, S.K. Seeley, Multidimensional Gas Chromatography: Fundamental Advances and New Applications, Anal. Chem. 85 (2013) 557–578. doi:10.1021/ac303195u.
[18] P.J. Marriott, S.-T. Chin, B. Maikhunthod, H.-G. Schmarr, S. Bieri, Multidimensional gas chromatography, TrAC, Trends Anal. Chem. 34 (2012) 1–21. doi:10.1016/j.trac.2011.10.013.
[19] M. Edwards, H. Boswell, T. Górecki, Comprehensive Multidimensional Chromatography, Current Chromatography. 2 (2015) 80–109. doi:10.2174/2213240602666150722232236.
[20] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, Journal of Chromatography A. 1255 (2012) 3–11. doi:10.1016/j.chroma.2012.05.050.
[21] Z. Zeng, J. Li, H.M. Hugel, G. Xu, P.J. Marriott, Interpretation of comprehensive two-dimensional gas chromatography data using advanced chemometrics, TrAC Trends in Analytical Chemistry. 53 (2014) 150–166. doi:10.1016/j.trac.2013.08.009.
[22] M.S. Klee, J. Cochran, M. Merrick, L.M. Blumberg, Evaluation of conditions of comprehensive two-dimensional gas chromatography that yield a near-theoretical maximum in peak capacity gain, Journal of Chromatography A. 1383 (2015) 151–159. doi:10.1016/j.chroma.2015.01.031.
[23] N.E. Watson, W.C. Siegler, J.C. Hoggard, R.E. Synovec, Comprehensive Three-Dimensional Gas Chromatography with Parallel Factor Analysis, Anal. Chem. 79 (2007) 8270–8280. doi:10.1021/ac070829x.
[24] W.C. Siegler, J.A. Crank, D.W. Armstrong, R.E. Synovec, Increasing selectivity in comprehensive three-dimensional gas chromatography via an ionic liquid stationary phase column in one dimension, Journal of Chromatography A. 1217 (2010) 3144–3149. doi:10.1016/j.chroma.2010.02.082.
[25] N.E. Watson, H.D. Bahaghighat, K. Cui, R.E. Synovec, Comprehensive Three-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry, Anal. Chem. 89 (2017) 1793–1800. doi:10.1021/acs.analchem.6b04112.
[26] N.E. Watson, S.E. Prebihalo, R.E. Synovec, Targeted analyte deconvolution and identification by four-way parallel factor analysis using three-dimensional gas chromatography with mass spectrometry data, Analytica Chimica Acta. 983 (2017) 67–75. doi:10.1016/j.aca.2017.06.017.
[27] C.E. Freye, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, Talanta. 161 (2016) 675–680. doi:10.1016/j.talanta.2016.09.002.
[28] A.M. Muscalu, M. Edwards, T. Górecki, E.J. Reiner, Evaluation of a single-stage consumable-free modulator for comprehensive two-dimensional gas chromatography: Analysis of polychlorinated biphenyls, organochlorine pesticides and chlorobenzenes, Journal of Chromatography A. 1391 (2015) 93–101. doi:10.1016/j.chroma.2015.02.074.
[29] J. Luong, X. Guan, S. Xu, R. Gras, R.A. Shellie, Thermal Independent Modulator for Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 88 (2016) 8428–8432. doi:10.1021/acs.analchem.6b02525.

[30] J.V. Seeley, N.E. Schimmel, S.K. Seeley, The multi-mode modulator: A versatile fluidic device for two-dimensional gas chromatography, Journal of Chromatography A. in press (2017). doi:10.1016/j.chroma.2017.06.030.
[31] R.B. Wilson, W.C. Siegler, J.C. Hoggard, B.D. Fitz, J.S. Nadeau, R.E. Synovec, Achieving high peak capacity production for gas chromatography and comprehensive two-dimensional gas chromatography by minimizing off-column peak broadening, Journal of Chromatography A. 1218 (2011) 3130–3139. doi:10.1016/j.chroma.2010.12.108.
[32] L.M. Blumberg, Accumulating resampling (modulation) in comprehensive two-dimensional capillary GC (GC×GC), J. Sep. Science. 31 (2008) 3358–3365. doi:10.1002/jssc.200800424.
[33] W. Khummueng, J. Harynuk, P.J. Marriott, Modulation Ratio in Comprehensive Two-dimensional Gas Chromatography, Anal. Chem. 78 (2006) 4578–4587. doi:10.1021/ac052270b.
[34] W.C. Siegler, B.D. Fitz, J.C. Hoggard, R.E. Synovec, Experimental Study of the Quantitative Precision for Valve-Based Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 83 (2011) 5190–5196. doi:10.1021/ac200302b.
[35] C.E. Freye, H.D. Bahaghighat, R.E. Synovec, Comprehensive two-dimensional gas chromatography using partial modulation via a pulsed flow valve with a short modulation period, Talanta. 177 (2018) 142–149. doi:10.1016/j.talanta.2017.08.095.
[36] C.E. Freye, L. Mu, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography, Journal of Chromatography A. 1424 (2015) 127–133. doi:10.1016/j.chroma.2015.10.098.
[37] C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, Partial least squares analysis of rocket propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection, Talanta. 153 (2016) 203–210. doi:10.1016/j.talanta.2016.03.016.
[38] R.B. Wilson, J.C. Hoggard, R.E. Synovec, High throughput analysis of atmospheric volatile organic compounds by thermal injection – isothermal gas chromatography – time-of-flight mass spectrometry, Talanta. 103 (2013) 95–102. doi:10.1016/j.talanta.2012.10.013.
[39] D.K. Pinkerton, B.A. Parsons, T.J. Anderson, R.E. Synovec, Trilinearity deviation ratio: A new metric for chemometric analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Analytica Chimica Acta. 871 (2015) 66–76. doi:10.1016/j.aca.2015.02.040.
[40] J.L. Hope, K.J. Johnson, M.A. Cavelti, B.J. Prazen, J.W. Grate, R.E. Synovec, High-speed gas chromatographic separations with diaphragm valve-based injection and chemometric analysis as a gas chromatographic "sensor," Analytica Chimica Acta. 490 (2003) 223–230. doi:10.1016/S0003-2670(03)00670-6.
[41] G.M. Gross, B.J. Prazen, J.W. Grate, R.E. Synovec, High-Speed Gas Chromatography Using Synchronized Dual-Valve Injection, Anal. Chem. 76 (2004) 3517–3524. doi:10.1021/ac049909g.

[42] H. Smith, R.D. Sacks, Column Selectivity Programming and Fast Temperature Programming for High-Speed GC Analysis of Purgeable Organic Compounds, Anal. Chem. 70 (1998) 4960–4966. doi:10.1021/ac980463b.
[43] T.-Y. Chen, M.-J. Li, J.-L. Wang, Sub-second thermal desorption of a micro-sorbent trap for the analysis of ambient volatile organic compounds, Journal of Chromatography A. 976 (2002) 39–45. doi:10.1016/S0021-9673(02)01073-7.

Chapter 6. Partial Least Squares Analysis of Rocket Propulsion Fuel Data Using Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography Coupled with Flame Ionization Detection

This chapter was reproduced from C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, "Partial Least Squares Analysis of Rocket Propulsion Fuel Data Using Diaphragm Valve-based Comprehensive Two-Dimensional Gas Chromatography Coupled with Flame Ionization Detection" *Talanta* 153 (2016) 203-210.

6.1 INTRODUCTION

Although aerospace fuel commodities are produced to meet numerous specification requirements, even advanced kerosene-based fuels such as rocket propellant (RP-1) may possess wide variation in chemical composition, which may result in unacceptable or unexpected performance and fuel properties [1–4]. Variation may be due to original crude oil/feed stock composition and refinery operating conditions. Additionally, RP-1 chemical composition can vary depending upon post-refinery formulation. Investigation of RP-1 fuel properties and their relationship to their chemical composition must be performed in order to develop optimum fuels for specific applications [5–11]. Although there are current methods to analyze RP-1 and related fuels [3,6,7,11], more informative, yet simpler and cost effective chemical analysis methods should be developed and investigated for routine application.

Gas chromatography (GC) coupled with flame ionization detection (GC – FID) is a traditional analytical platform to separate and analyze volatile and semi-volatile mixtures [12]. However, complex mixtures such as RP-1 fuels generally cannot be sufficiently separated with a single dimension of GC in order to provide the desired analytical insight. Comprehensive two-

dimensional (2D) gas chromatography with flame ionization detection (GC × GC – FID) can improve upon the separation power of one dimensional GC and yield quantitative information [13–15]. With GC × GC, two "orthogonal" separation dimensions are implemented to increase peak capacity which substantially improves the information that may be gleaned from complex samples. Typically, the first GC × GC separation dimension (^{1}D) uses a column with a non-polar stationary phase, while the second dimension (^{2}D) uses a column with a polar stationary phase, providing complementary information. However, a "reversed column" GC × GC configuration has been found to be more appropriate to provide optimized use of the 2D separation space to effectively separate the various compound classes in RP-1 and related fuels (*n*-alkanes, branched alkanes, mono- di- and tricyclic alkanes, and aromatics) [16,17]. The reversed column configuration uses a polar ^{1}D column followed by a non-polar ^{2}D column [18].

The modulator is the heart of the GC × GC instrument, interfacing the ^{1}D column to the ^{2}D column to accomplish a comprehensive two-dimensional separation. Here, we categorize three modulator designs: cryogenic-based "thermal" modulation [19–23], flow modulation [24–26], and diaphragm valve modulation [14,27,28]. While thermal modulation is popular, it can be expensive to implement and is generally not amenable to reduce to a simple platform because of the complexity involved with the cryogenic fluids. Flow modulation is becoming more popular with GC × GC, however, the injected pulse width onto the ^{2}D column is intrinsically quite large, resulting in a significant source of peak broadening and a relatively low ^{2}D peak capacity [29,30]. In comparison, diaphragm valve modulation is relatively simple and inexpensive to implement, while also providing a narrow injected pulse width onto the ^{2}D column with superior peak capacity even when a short modulation period is used (e.g., 1–2 s) [14,27,28,31–33]. The primary shortcoming of the diaphragm valve in the past has been the limited temperature range due to the

O-rings within the valve, limiting the usage to a maximum of 175 °C if placed directly in the GC oven. This shortcoming was overcome to a reasonable extent by face mounting the valve on the top of the GC case to preserve the temperature sensitive O-rings within the valve, with a maximum temperature of 265 °C [28]. Recently, a state-of-the-art "high temperature" diaphragm valve with perfluoroelastomer O-rings has been developed that is capable of being placed directly in the GC oven; it can withstand temperatures up to 325 °C [34]. Recently, we reported the implementation and evaluation of this high temperature diaphragm valve with GC × GC – FID [35]. This high temperature diaphragm valve-based GC × GC – FID is implemented in this current study.

GC × GC – FID is well suited technique for when identification and quantification of every chemical compound is generally not possible and/or is an extremely laborious and impractical process for complex samples such as RP-1 fuels. Chromatographic-based ASTM methods often are rooted in this type of strategy. In this regard, it is advantageous to implement chemometric analysis of complex samples, specifically herein to discover important chemical/physical relationships between compositionally unique RP-1 fuels [16,17,36]. Partial least squares (PLS) analysis is well suited to quantitatively relate the GC × GC – FID data of complex samples, such as RP-1 fuels, with other forms of chemical/physical measurements.

PLS analysis provides the following two valuable outcomes. First, using a training set of samples, a linear correspondence of the chemical/physical properties can be modeled using the GC × GC – FID data, so subsequent analyses of GC × GC – FID data of new samples can be used to predict the modeled chemical/physical property without having to directly measure these properties on the new samples. Second, the underlying relationship between the chemical composition of the samples can be correlated to the modeled chemical/physical measurements, so a deeper understanding between chemical composition via the GC × GC – FID data and the

chemical/physical measurements is provided [37–40]. To achieve both of these outcomes, leave-one-out cross validation is often performed (LOOCV) [16,17,37].

In this study, we seek to demonstrate the simple, yet highly capable GC × GC – FID instrumental platform, combined with PLS, to provide valuable information about the chemical composition of RP-1 fuels. Due to the potential for significant variation in chemical composition of RP-1 fuels as a result of variations that may result during the distillation and blending processes, it is often beneficial to evaluate special laboratory blends where the analyst can control the make-up of the fuel [5–9]. Therefore, we study a set of fuels that have been well studied before, so as to provide validation of the instrument performance. Furthermore, we present implementation of a state-of-the-art high temperature diaphragm valve [34,35] as the modulator, with the valve mounted directly in the oven. To the best of our knowledge, this is the first application study using the high temperature diaphragm valve. With GC × GC – FID, we hope to obtain similar quantitative information as previously obtained while using GC × GC – TOFMS [16]. While GC × GC – TOFMS can yield substantially more qualitative information especially with the ability to identify analytes, instrumental cost is too high for routine implementation in many laboratory settings. Furthermore, GC × GC – TOFMS requires a large amount of upkeep, such as the systems associated with maintaining the vacuum, to keep it operational. When identification of analytes is not necessary or has already been performed, a FID can provide ample information and requires less maintenance and lower operational costs.

The following studies were performed. First, the summed GC × GC – FID signal of three compound-class selective 2D regions (alkanes, cycloalkanes, and aromatics) was simply regressed against previously measured (ASTM standard test method) values for these compound classes, and LOOCV was used for regression validation. Next, for comparison, using PLS analysis with

LOOCV, the GC × GC – FID signal of the entire 2D separations was regressed against the same ASTM values for the same three compound classes. Using PLS, retention times are preserved and it is possible to determine how specific compounds influence the model by looking at the linear regression vectors (LRVs). Finally, PLS analysis was used to provide a more detailed study of compound classes (*n*-alkanes, *iso*-alkanes, mono-,di-,and tri-cycloalkanes, and aromatics) and of physical properties previously determined by ASTM methods (such as net heat of combustion, hydrogen content, density, kinematic viscosity, sustained boiling temperature and vapor rise temperature). Results from these studies using GC × GC – FID instrumentation are compared to previously reported results using GC × GC – TOFMS [16].

6.2 EXPERIMENTAL

Ten previously studied RP-1 fuel samples (listed in Table 6.1) were obtained from the Air Force Research Laboratory (AFRL) [5,6]. HPLC Grade hexane and acetone were obtained from Fisher Scientific and were used as solvent rinses prior to injection. Chromatographic data was obtained using a GC × GC – FID consisting of an Agilent 6890 GC (Agilent Technologies, Palo Alto, CA, USA) modified in-house with a high-speed, six-port diaphragm valve VICI model DV-12-1116T (Valco Instruments Company Inc., Houston, TX, USA) fitted with a 10 µl sample loop, with the complete valve mounted inside the GC oven. This valve allows separation temperatures up to 325 °C without degradation of the valve's temperature sensitive O-rings [34]. The diaphragm valve actuation controlled using a LabVIEW 8.2 program (National Instruments, Austin, TX, USA). The stock electrometer for the Agilent FID was replaced with a high-speed electrometer built in-house allowing the data to be collected at 100 kHz, with the data further boxcar averaged to 1 kHz, using an FID temperature of 250 °C. The primary column (^{1}D) for the GC × GC used a RTX-wax stationary phase: 30 m length, 250 µm inner diameter, and 0.5 µm film thickness. The

secondary column (^{2}D) used a RTX-1 stationary phase: 1.2 m length, 100 μm inner diameter, and 0.18 μm film thickness. The modulation period was set to 2.5 s (i.e., the secondary column separation time) with a pulse width of 200 ms which was sufficient to clear the sample loop. The RP-1 fuel samples were introduced to the GC × GC via a 7683B auto-injector (Agilent Technologies, Palo Alto, CA, USA). Prior to injection, hexane and acetone were used as solvent rinses. One μl of the RP-1 fuels was injected via the auto-injector, using a 100:1 split, and an inlet temperature of 225 °C. Flow on the first column was 1.8 ml/min while flow on the second column was 2.5 ml/min. Oven temperature was held at 40 °C for 2 min, increased at 6 °C/min to 225 °C and held for 3 min. The secondary column oven temperature control was not applied, so the secondary column oven was open and at the same nominal temperature as the primary column oven.

RP-1 sample	NIST number	AFRL designation
1	11	LB080409-01
2	10	LB073009-06
3	9	LB073009-08
4	8	LB080409-05
5	7	LB073009-05
6	5	LB073009-01
7	4	LB073009-09
8	1	LB073009-02
9	2	LB073009-03
10	3	XC2521HW10

Table 6.1 RP-1 fuel set. The NIST and ARFL numbers are provided for reference to previous studies [5,6].

For each RP-1 sample, four replicates of GC × GC – FID chromatographic data were collected. The chromatographic data was imported into MATLAB2012b (MathWorks, Natick, MA) where it was boxcar smoothed by a factor of 3 and baseline corrected using in-house software.

Data was binned 7.5 s along ^{1}D (3 modulations) and 10 ms along ^{2}D. Binning was performed to remove any chromatographic misalignment and reduce the size of the data set to save computation time [41–43]. The data was then forwarded to PLS Toolbox 7.5 (Eigenvector Research Inc., Wenatchee, WA, USA) and underwent mean centering [36]. The ASTM measured data was imported to PLS Toolbox and underwent autoscaling. The following compound class analyses (mass%) were performed using ASTM D2425: paraffins (*n*-and *iso*-), cycloparaffins, dicycloparaffins, and tricycloparaffins. Within the results reported herein, the term "alkanes" is used in place of "paraffins." *n*-alkane analysis provided further categorization; *iso*-alkane content was assumed by subtraction. Total aromatic content (mass%) was determined according to ASTM D6379. Additional composition and physical properties of these fuels have been reported [5,6,44]. The following list includes the physical properties and method of measurement: density (g/ml) – ASTM D4052 (at 15 °C), kinematic viscosity (mm^2/s) – ASTM D445 (at −10 °C), net heat of combustion (MJ/kg) – ASTM D4809, net heat of combustion (MJ/l) – ASTM D4809 (density at 30 °C), hydrogen content (wt%) – Perkin Elmer Elemental Analyzer, Model EA2400, sustained boiling temperature (°C), and vapor rise temperature (°C).

For the PLS analysis of GC × GC – FID data, two data matrices (X and Y block) are related through calculation of loadings referred to as latent variables (LVs). In our study, the X block is the chromatographic data obtained via GC × GC – FID analysis, while the Y block is the measured values [5,6,44]. PLS aims to model the covariance between these two matrices by finding the multidimensional direction in the X-block which explains the maximum multidimensional variance in the Y-block. In other words, PLS uses analytes which have a large range of intensities across the different RP-1 fuels and attempts to correlate them to the differences in the measured values. PLS analysis seeks to provide a linear trend relating the chromatographic data to the

chemical/physical properties measured via ASTM methods. For the construction of the PLS models, all underwent LOOCV for testing the overall predictive ability of each model, and to determine the most appropriate number of LVs. The LOOCV procedure for N samples uses $N-1$ samples to predict the omitted sample. This is repeated N times and the model is analyzed yielding a root mean square error of cross validation (RMSECV) [45]. The RMSECV values indicate the precision, and to some extent the accuracy, of the PLS models through the calculated residuals of the model. The coefficient of determination, R^2, is another manner in which to calculate the goodness-of-fit, but RMSECV values are more appropriate herein because RMSECV explicitly indicates to what extent the predictions deviate. R^2 values are mentioned but excluded from Tables for sake of brevity. Normalized root mean square error of cross validation (NRMSECV) is found by dividing the RMSECV by the range of the measured values. This value is dimensionless and is useful for comparing models with different units. Additionally, LRVs can be examined to determine which analyte compounds are positively correlated (or negatively correlated) with the specific property of interest [16]. Analytes that exhibit positive values in the LRVs are correlated with increasing the specified property while negatively correlated analytes will exhibit negative values in the LRV and are correlated with decreasing the specified property.

6.3 RESULTS AND DISCUSSION

Using GC × GC – FID in reversed column configuration, we achieved an excellent 2D chromatographic separation of the compound classes: alkanes (*n*-and *iso*-), cycloalkanes, and aromatics in the RP-1 fuels. Fig. 6.1 shows the 2D separation of a representative fuel LB073009-08. The chromatograms for the other nine RP-1 fuels are excluded for brevity. For the analysis of the RP-1 fuels, regions of the chromatogram were chosen based on previous work [16] and expected location of the compound classes within the chromatogram. As seen in the figure, the

alkane compounds are located from 2.0 to 2.5 s and 0.0 to 0.4 s on ^{2}D. The alkanes were purposely allowed to wrap around to efficiently use the 2D separation space. The cycloalkanes were located from 0.9 to 2.0 s on ^{2}D. The aromatics are located from 15 to 36 min on ^{1}D, and from 0.4 to 0.9 s on ^{2}D. Although it is possible to discern subclasses for the cycloalkanes and aromatic classes (i.e. monocyclics, dicyclics, tricyclics, mono-aromatics, and di-aromatics), it would be extremely difficult to pick an exact boundary especially for the cyclic compounds without knowing the identity of the chromatographic peaks. However, the emphasis of the present study is not exhaustive peak identification, but rather demonstrating the ability to qualitatively and quantitatively relate the RP-1 fuels based on PLS modeling and LOOCV of the chromatographic data in comparison to the ASTM derived values.

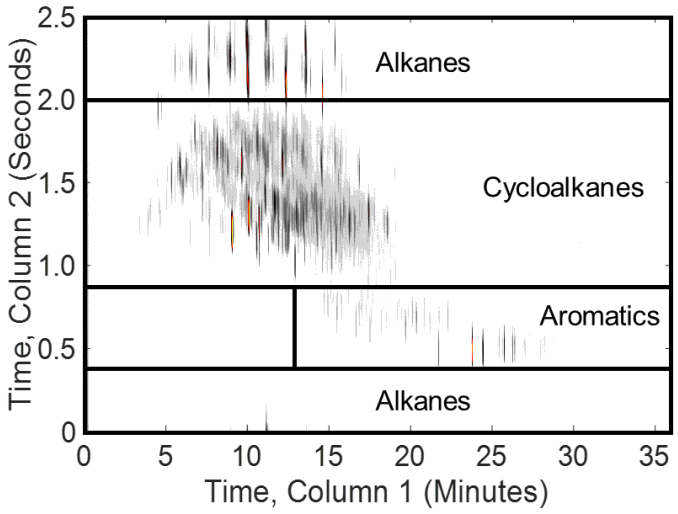


Figure 6.1 A representative GC × GC – FID chromatogram for the RP-1 fuels (LB073009- 08). The alkane region encompassed the entire ^{1}D separation, and 2.0–2.5 s and 0– 0.4 s on ^{2}D. The cycloalkanes encompassed the entire ^{1}D separation, and from 0.9 to 2.0 s on ^{2}D. Finally, the aromatics were located in the area from 13 min to the end of the ^{1}D separation, and from 0.4 to 0.9 s on ^{2}D.

Measured Property	Range	RMSECV	NRMSECV
Alkanes (% Mass)	39.9-52.6	0.855±0.067	6.73%
Cycloalkanes (% Mass)	43.8-59.9	0.734±0.008	4.56%
Aromatics (% Mass)	0.2–3.7	0.530±.290	15.1%

Table 6.2 Summary of the LOOCV regression model for the measured properties analyzed via GC × GC – FID. RMSECV values for each class of compounds are included as well as the NRMSECV.

To determine the percent composition for each of the principal hydrocarbon classes (alkanes, cycloalkanes, aromatics), the signal in the specified area was summed and divided by the total signal in the chromatogram, yielding a percent signal for each compound class. These calculations were performed for the ten RP-1 fuels. The values obtained by the ASTM D2425 method were summed where appropriate to yield the principal hydrocarbon classes (monocycloalkanes, dicycloalkanes, and tricycloalkanes were summed to yield cycloalkanes). Using LOOCV, the results from the GC × GC – FID method were regressed against the measured mass% values. This was accomplished by regressing the summed FID signal in the region of interest against the measured values. Table 6.2 shows the RMSECV and NRMSECV values of the three classes. The regression using this GC × GC – FID method provides high accuracy and precision for predicting the ASTM values for alkanes and cycloalkanes; however, this method is somewhat less precise for the prediction of aromatics, potentially due to the lower signal-to-noise ratio of the aromatic peaks (since they are present in the fuels at a relatively low concentration). LOOCV yielded a linear trend between the predicted and measured values. Fig. 6.2A shows the LOOCV regression plots for the GC × GC – FID alkane class signal percentage method. Likewise, the LOOCV regression plots for the cycloalkanes and aromatics are shown in Fig. 6.2B and C.

Error bars (± one standard deviation) are included for these figures and subsequent figures based on the four injection replicates.

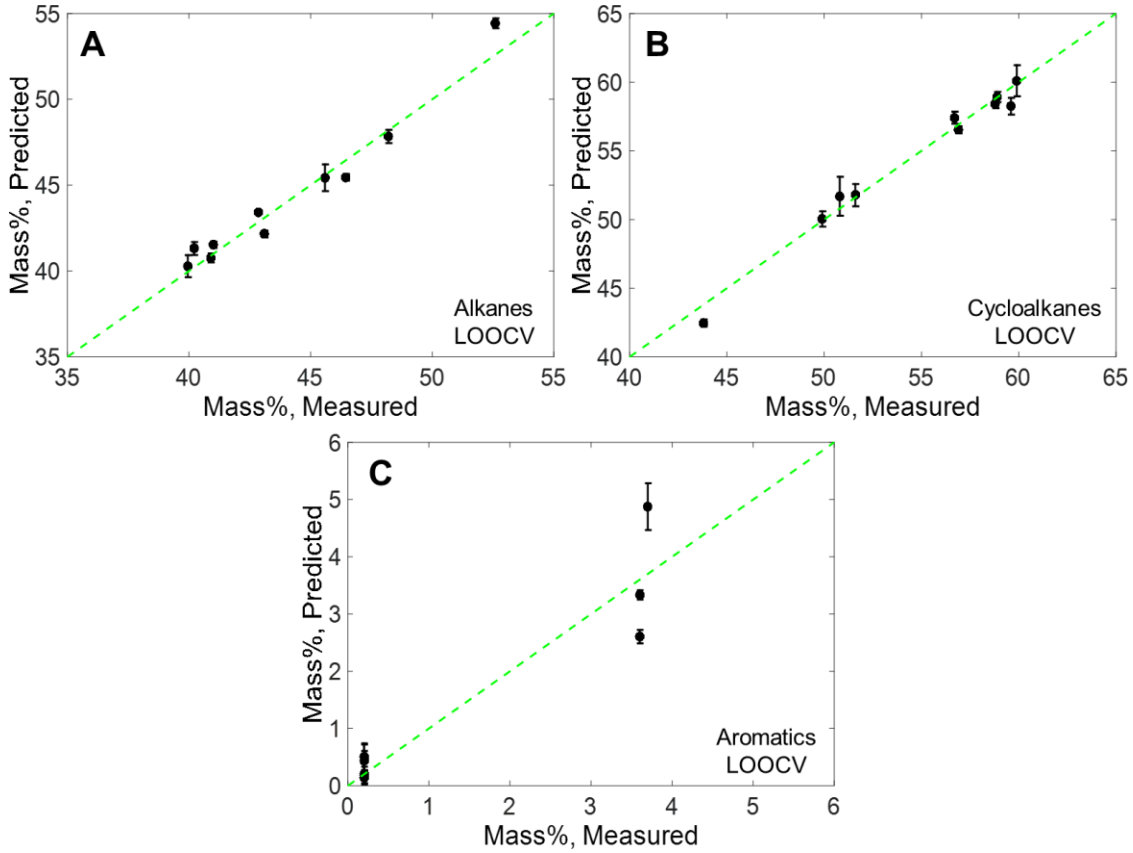


Figure 6.2 Results for the simple regression using LOOCV for the compound classes derived from the GC × GC – FID chromatograms. After preprocessing (see Experimental), the total area of the chromatogram where the respective compounds were located was summed. This value was then regressed against the ASTM derived values. The dashed line shows the ideal agreement between the predicted and measured mass%. (A) LOOCV of alkanes. (B) LOOCV of cycloalkanes. (C) LOOCV of aromatics.

Next, the GC × GC – FID data was analyzed using PLS with the previously measured mass% content for the principal hydrocarbon classes: alkanes, cycloalkanes, and aromatics. Chromatographic data from the appropriate section of the chromatogram was forwarded to PLS

which was regressed against the measured values. The PLS model for the alkanes, cycloalkanes, and aromatics are shown in Fig. 6.3A–C, respectively. A summary of the RMSECV and NRMSECV values is provided in Table 6.3.

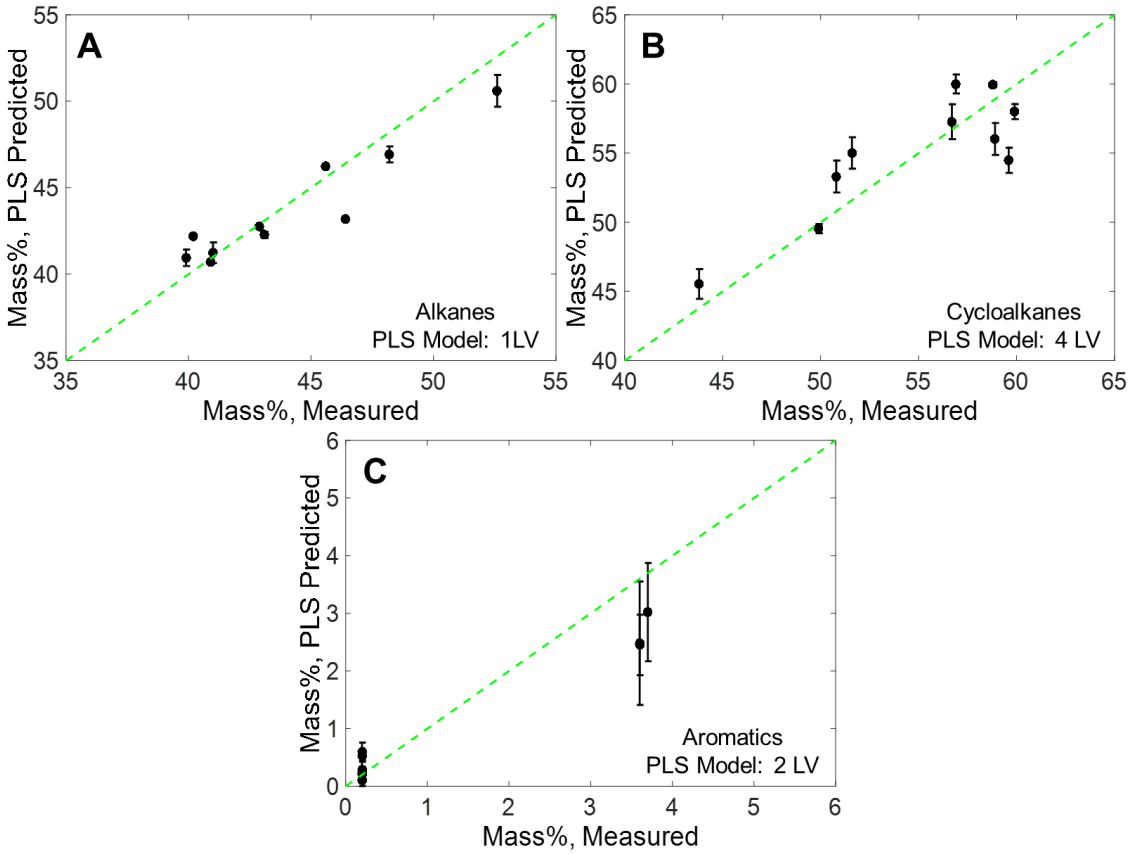


Figure 6.3 PLS modeling of the GC × GC – FID chromatograms via LOOCV for the compound classes. The dashed line shows the ideal agreement between the predicted and measured mass%. (A) PLS modeling of alkanes using 1 LV. (B) PLS modeling of cycloalkanes using 4 LVs. (C) PLS modeling of aromatics using 2 LVs.

The simple regression using LOOCV yields a more linear trend than the PLS analysis, which results in significantly lower RMSECV and NRMSECV. The simple regression method using LOOCV most likely performs better since analytes in the same class (i.e. alkanes, cycloalkanes, and aromatics) have almost identical FID responses. However, it should be noted

that the advantage of PLS compared to the simple regression is the preservation of the chromatographic information. The number of LVs was selected after inspecting the RSMECV values. When there is an insignificant difference between two satisfactory PLS models, the rule of parsimony is adhered to and the model with fewer LVs is chosen [46]. For the alkanes, the PLS model was built using 1 LV as seen in Fig. 6.3A. The PLS predicted mass% is plotted versus the measured mass% as conventionally shown. Similarly, the PLS model for the cycloalkanes was constructed with 4 LV, and the aromatics PLS modeling was constructed with 1 LV. Comparison of predicted and measured values for these compound classes is shown in Fig. 6.3B and C, respectively.

Measured Property	Range	RMSECV	NRMSECV
Alkanes (% Mass)	39.9-52.6	1.52±0.11	12.0%
Cycloalkanes (% Mass)	43.8-59.9	2.76±0.25	17.1%
Aromatics (% Mass)	0.2–3.7	0.945±0.18	27.0%

Table 6.3 Summary of the PLS model metrics using LOOCV for the measured properties analyzed via GC × GC – FID. RMSECV values for each class of compounds are included as well as the NRMSECV.

In the third portion of the study, the mass% of *n*-alkanes, *iso*-alkanes, monocycloalkanes, dicycloalkanes, tricycloalkanes, and aromatics was modeled via PLS. The PLS models were constructed using only the relevant portions of the chromatographic regions as previously shown in Fig. 6.1 after the data was binned (e.g. *n*-alkanes and *iso*-alkanes were predicted using only the alkane region). A summary of the RMSECV and NRMSECV values is provided in Table 6.4. Included within the table are previous results using GC × GC – TOFMS [16]. The GC × GC – FID data with PLS modeling yields very similar results when compared to using GC × GC – TOFMS

for *n*-alkanes, *iso*-alkanes, and monocycloalkanes. GC × GC – FID struggles to some extent when predicting the di- and tricycloalkanes and the aromatics. The NRMSECV values for the tricycloalkanes are relatively high. This is most likely due to the small sample amount in each sample resulting in low intensities. Furthermore, the tricycloalkanes have an extremely small range, which results in a very challenging situation for PLS modeling, causing the NRMSECV values to be higher in comparison to the other studied chemical compositions. A potential reason that aromatics respective RMSECV and NRMSECV values are significantly higher for the GC × GC – FID method is the low intensities of these peaks. Another possible reason for the high RMSECV values for the different classes could be the potential for inherent inaccuracy of the ASTM methods. Measurements of the ASTM values are very reproducible with variations of less than 3%, but method accuracy has not been determined because there is no suitable reference [47]. ASTM methods are inferred as the gold standard for providing the correct (accurate) answer, however there are obvious shortcomings to this common perception.

		GC×GC-FID		GC×GC-TOFMS	
Measured Property	**Range**	**RMSECV**	**NRMSECV**	**RMSECV**	**NRMSECV**
n-Alkane (% Mass)	0.5-14	1.14±0.29	8.44%	1.34±0.10	9.93%
Iso-Alkane (% Mass)	36.2–41.5	1.04±0.06	19.6%	1.14±0.04	21.4%
Monocycloalkanes (% Mass)	26.5–38.0	2.15±0.13	18.7%	2.46±0.28	21.3%
Dicycloalkanes (% Mass)	13.5–18.8	1.53±0.093	28.9%	0.84±0.21	15.8%
Tricycloalkanes (% Mass)	3.50–4.80	0.453±0.024	34.8%	0.25±0.04	19.2%
Aromatics (% Mass)	0.2–3.7	0.699±0.202	19.8%	0.075±0.007	2.14%

Table 6.4 Summary of the PLS model metrics using LOOCV for the measured properties analyzed. RMSECV values for the GC × GC – FID and GC × GC – TOFMS [16] and their accompanying normalized RMSECV.

Resulting from this third portion of the study, the PLS regression plots for the *n*-alkanes, monocycloalkanes, dicycloalkanes, and tricycloalkanes are shown in Fig. 6.4A–D, respectively. The other PLS regressions are not included for brevity. Modeling the specific subclasses such as the mono-, di-, and tricycloalkanes can be performed to obtain more information about them. However, these subclasses can be harder to predict because they span a very small mass% range.

As seen in Fig. 6.4A, the *n*-alkanes PLS modeling was performed with 3 LV, and the regression exhibits a very linear trend with a low amount of variation. The *n*-alkane PLS regression had a R^2 value of 0.933, a RMSECV value of 1.14 mass%, and a NRMSECV of 8.44% indicating a high precision (see Table 6.4). A large number of samples lay between 0.5% and 3% as seen in the PLS regression. The monocycloalkanes in Fig. 6.4B were modeled using 5 LV, the dicycloalkanes in Fig. 6.4C were modeled using 2 LV, and the tricycloalkanes in Fig. 6.4D were modeled using 1 LV. The monocycloalkanes had a R^2, RMSECV, and NRMSECV values of 0.694, 2.15 mass%, and 18.7%, respectively; the dicycloalkanes had a R^2, RMSECV, and NRMSECV values of 0.246, 1.53 mass%, and 28.9%, respectively; and the tricycloalkanes had R^2, RMSECV, and NRMSECV values of 0.062, 0.453 mass%, and 34.8%, respectively (Table 6.4). The monocycloalkanes had the largest mass% range of all the RP-1 fuels studied. Most of the values lie close to the line with a slope of unity, which indicates good agreement between predicted and measured value; some values in the lower range are slightly over/under predicted. Most of the dicycloalkanes seem to be over predicted especially near the middle of the mass% range. The tricycloalkanes are over predicted at the lower values and under predicted at the higher values. The tricycloalkanes span a very small mass% range, making this a difficult compositional trait to accurately predict, as is also suggested by the low extremely R^2 value of 0.062. In summary, it has been shown that GC × GC – FID chromatographic data can be readily correlated with the measured

chemical compositional values via PLS modeling, in some cases better than others. From PLS modeling of the chemical content based on measured values, LRVs are obtained. In basic terms, the LRVs describe what analytes are positively correlated and negatively correlated with the measured mass% values. Although not specifically presented herein, LRVs could be used to determine which analytes affect the mass% in the modeling.

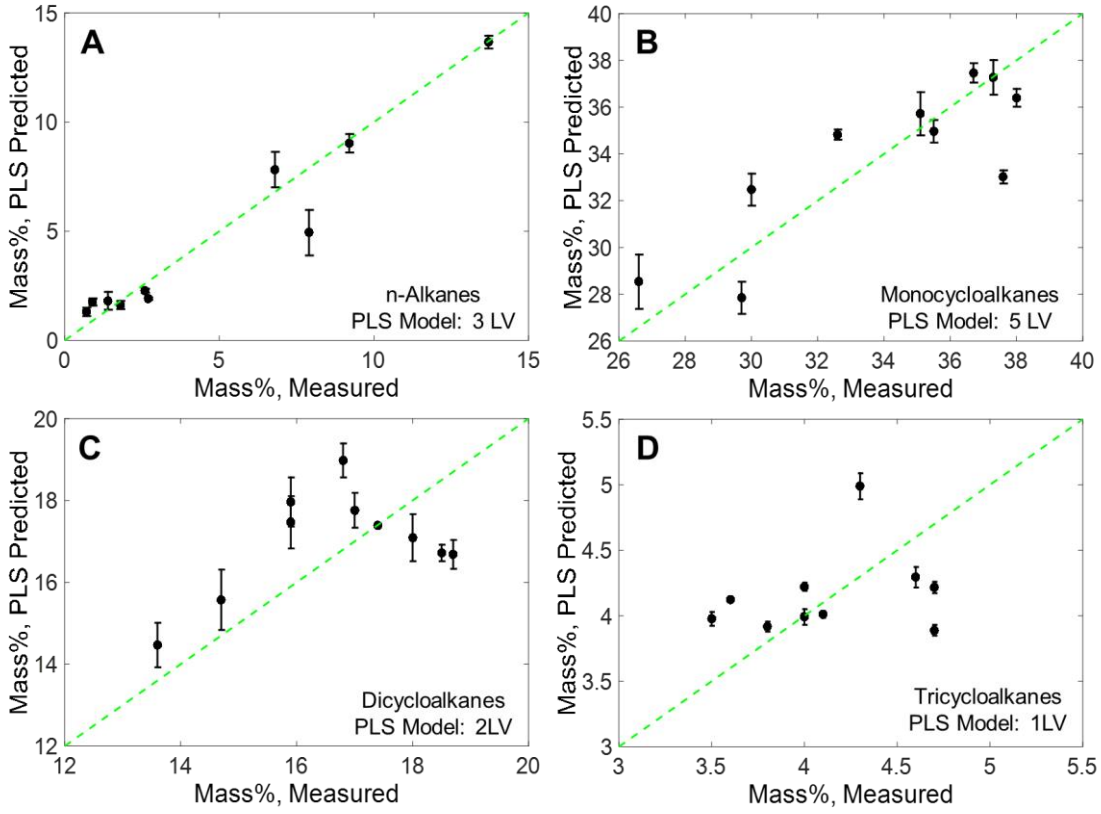


Figure 6.4 PLS modeling of the GC × GC – FID chromatograms via LOOCV for the compound classes. (A) PLS modeling of n-alkanes using 3 LVs. (B) PLS modeling of monocyloalkanes using 2 LVs (C) PLS modeling of dicycloalkanes using 2 LVs. (D) PLS modeling of tricycloalkanes using 1 LV.

Finally, physical properties obtained using standard ASTM test methods were studied with PLS; a summary of the RMSECV and NRMSECV values for the GC × GC – FID and previous work using GC × GC – TOFMS are provided in Table 6.5. The GC × GC – FID RMSECV values are very similar to the GC × GC – TOFMS RMSECV values, with most of them being within a few percent of each other. Fig. 6.5 shows the PLS model results of the net heat of combustion (MJ/l), which is representative of the other physical properties (not shown). In the PLS regression of the net heat of combustion as provided in Fig. 6.5, the point at the low extreme of the data set is over predicted. This is mainly due to the use of LOOCV with the PLS modeling which tends to amplify slight deviations from the trend at extremes in the data sets.

		GC×GC-FID		GC×GC-TOFMS	
Measured Property	**Range**	**RMSECV**	**NRMSECV**	**RMSECV**	**NRMSECV**
Hydrogen content (% wt)	14.15-14.45	0.0623±0.0029	20.8%	0.055±0.007	18.3%
Density(g/L)	805.0-816.0	2.03±0.07	18.5%	1.9±0.03	16.9%
Kinematic Viscosity(mm^2/s)	4.80-5.70	0.204±0.013	22.7%	0.13±0.02	14.8%
Net heat of combustion (MJ/L)	34.4-34.9	0.0671±0.0007	13.4%	0.065±0.001	13.0%
Neat heat of combustion (MJ/kg)	43.25-43.45	0.0303±0.0024	15.2%	0.42±0.004	20.8%
Sustained boiling temp. (°C)	204.5-212.5	1.53±0.14	19.1%	1.4±0.01	17.4%
Vapor rise temp. (°C)	207-214	1.70±0.11	24.3%	1.4±0.01	20.3%

Table 6.5 Summary of the PLS model metrics using LOOCV for the measured properties analyzed. RMSECV values for each physical property are included for GC × GC – FID and GC × GC – TOFMS [16] and their accompanying normalized RMSECV.

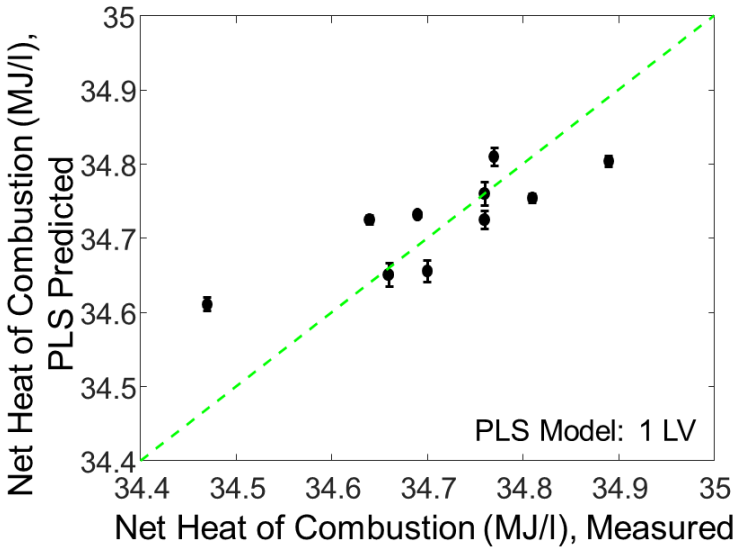


Figure 6.5 PLS regression using LOOCV for the net heat of combustion (MJ/l) using 1 LV. The dashed line indicates the ideal line.

Using the LRVs, specific analytes can be identified that directly affect the physical properties of the RP-1 fuels. Because there are thousands of chemical compounds, it can be extremely difficult to identify a compound although it can be accomplished based on ^{1}D and ^{2}D retention times. However, general trends can be seen in the LRVs such as those observed in Fig. 6.6A and B. The LRV for the net heat of combustion (MJ/l) indicates that the alkane compounds (which are located from 2.0 to 2.5 s on ^{2}D) and aromatics (which are located from 0.4 to 0.9 s on ^{2}D) are negatively correlated with the net heat of combustion while the monocycloalkanes and dicycloalkanes are positively correlated. It can be seen in the LRV that the dicycloalkanes are the most positively correlated with the net heat of combustion. By using LRVs the specific physical properties of the RP-1 fuel can be tailored for the specific fuel application need.

6.4 CONCLUSION

We have demonstrated the evaluation of RP-1 fuels using GC × GC – FID followed with PLS analysis with LOOCV. The GC × GC – FID instrumental platform was facilitated by a

recently developed state-of-the-art diaphragm valve (suitable for continuous high temperature operation) that was implemented directly in the oven as the GC × GC modulator. Both chemical composition and physical properties of these kerosene-based fuels were successfully analyzed. This method could be used to analyze a myriad of other complex volatile mixtures. Based on the information obtained in the physical properties modeling, specific properties of the fuel can be tailored for the specific purpose of the fuel. Overall, this technique should provide another viable, simple, inexpensive method for the analysis of kerosene-based rocket fuels.

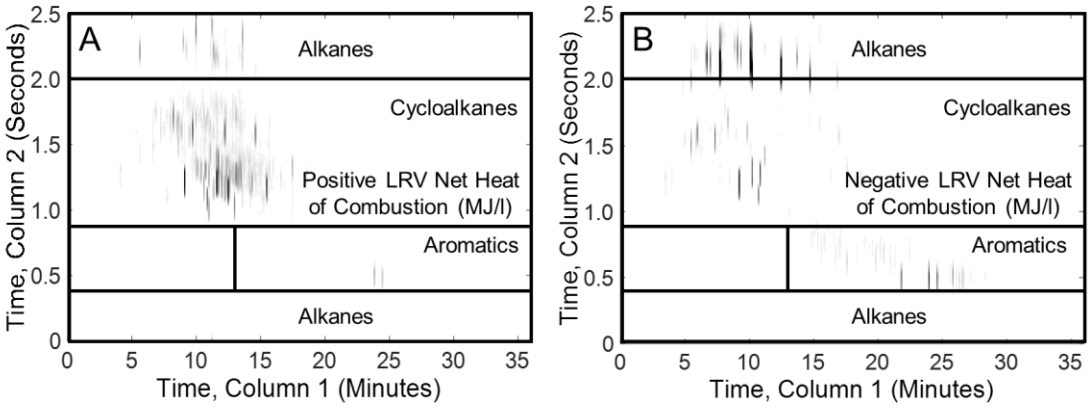


Figure 6.6 Linear regression vectors (LVRs) for the PLS modeling net heat of combustion (MJ/l). (A) Positive linear regression vector components. (B) Negative linear regression vector components.

6.5 ACKNOWLEDGEMENTS

The work at the University of Washington was performed under subcontract to ERC, Incorporated, Air Force Research Laboratory, Edwards AFB, CA. The RP-1 fuels were provided by the Air Force Research Laboratory/RQRC, Edwards AFB, CA. Compound class (hydrocarbon type) analysis was performed by Dr. Linda Shafer of the University of Dayton Research Institute (UDRI), Wright-Patterson AFB, OH. We would also like to thank Stan Stearns of Valco for the

high temperature diaphragm valves. Certain commercial equipment, instruments or materials are identified in this paper in order to adequately specify the experimental procedure. Such identification does not imply recommendation or endorsement by the University of Washington, the United States Air Force, or the National Institute of Standards and Technology, nor does it imply the materials or equipment identified are necessarily the best available for that purpose.

6.6 REFERENCES

[1] D. Cookson, B. Smith, Calculation of jet and diesel fuel properties using carbon-13 NMR spectroscopy, Energy Fuels 24 (1990) 152-156.

[2] G. Liu, L. Wang, H. Qu, H. Shen, X. Zhang, S. Zhang, Z. Mi, Artificial neural network approaches on composition–property relationships of jet fuels based on GC–MS,

[3] Fuel 86 (2007) 2551-2559.

[4] M.L. Huber, E.W. Lemmon, T.J. Bruno, Effect of RP-1 compositional variability on thermophysical properties, Energy Fuels 23 (2009) 5550-5555.

[5] M.J. DeWitt, T. Edwards, L. Shafer, D. Brooks, R. Striebich, S.P. Bagley, M.J. Wornant, Effect of aviation fuel type on pyrolytic reactivity and deposition propensity under supercritical conditions, Ind. Eng. Chem. Res. 50 (2011) 10434-10451.

[6] M.C. Billingsley, J.T. Edwards, L.M. Shafer, T.J. Bruno AIAA 2010-6824, 46th American Institute of Aeronautics and Astronautics (AIAA) Joint Propulsion Conference and Exhibit, Nashville, TN, July (2010)

[7] T.M. Lovestead, B.C. Windom, J.R. Riggs, C. Nickell, T.J. Bruno, Assessment of the compositional variability of RP-1 and RP-2 with the advanced distillation curve approach, Energy Fuels 24 (2010) 5611-5623.

[8] R.V. Gough, T.J. Bruno, Composition-explicit distillation curves of alternative turbine fuels, Energy Fuels 27 (2013) 294-302.

[9] P.Y. Hsieh, K.R. Abel, T.J. Bruno, Analysis of marine diesel fuel with the advanced distillation curve method, Energy Fuels 27 (2013) 804-810.

[10] J.L. Burger, T.J. Bruno, Application of the advanced distillation curve method to the variability of jet fuels, Energy Fuels 26 (2012) 3661-3671.

[11] N.J. Begue, J.A. Cramer, C. Von Bargen, K.M. Myers, K.J. Johnson, R.E. Morris, Automated method for determining hydrocarbon distributions in mobility fuels, Energy Fuels 25 (2011) 1617-1623.

[12] T.J. Bruno, L.S. Ott, T.M. Lovestead, M.L. Huber, The composition-explicit distillation curve technique: Relating chemical analysis and physical properties of complex fluids, J. Chromatogr. A 1217 (2010) 2703 (Extraction 2009 Special Issue – Invited Review)

[13] I.G. McWilliam, R.A. Dewar, Flame ionization detector for gas chromatography, Nature 181 (1958) 760.

[14] Z. Liu, J.B. Phillips, Comprehensive two-dimensional gas chromatography using an on-column thermal modulator interface, J. Chromatogr. Sci. 26 (1991) 227-231.

[15] C.A. Bruckner, B.J. Prazen, R.E. Synovec, Comprehensive two-dimensional high-speed gas chromatography with chemometric analysis, Anal. Chem. 70 (1998) 2796-2804.

[16] W. C. Siegler, B. D. Fitz, J. C. Hoggard, R. E. Synovec, Experimental study of the quantitative precision for valve-based comprehensive two-dimensional gas chromatography, Anal. Chem. 83 (2011) 5190-5196.
[17] B. Kehimkar, J.C. Hoggard, L.C. Marney, M.C. Billingsley, C.G. Fraga, T.J. Bruno, R.E. Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, J. Chromatogr. A 1327 (2014) 132-140.
[18] B. Kehimkar, B. A. Parsons, J. C. Hoggard, M. C. Billingsley, T. J. Bruno, R. E. Synovec, Modeling RP-1 fuel advanced distillation data using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry and partial least squares analysis, Anal. Bioanal. Chem. 407 (2015) 321-330.
[19] B. Omais, M. Courtiade, N. Charon, D. Thiébaut, A. Quignard, M.C. Hennion, Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate, J. Chromatogr. A 1218 (2011) 3233-3240.
[20] J.B. Phillips, R.B. Gaines, J. Blomberg, F.W.M. van der Wielen, J.M. Dimandja, V. Green, J. Granger, D. Patterson, L. Racovalis, H.J. de Geus, J. de Boer, P. Haglund, J. Lipsky, V. Sinha, E.B. Ledford Jr., A robust thermal modulator for comprehensive two-dimensional gas chromatography, J. High Resolut. Chromatogr. 22 (1999) 3-10.
[21] J.B. Phillips, J. Xu, Comprehensive multi-dimensional gas chromatography, J. Chromatogr. A 703 (1995) 327-334.
[22] P.J. Marriott, R.M. Kinghorn, Longitudinally modulated cryogenic system. A generally applicable approach to solute trapping and mobilization in gas chromatography, Anal. Chem. 69 (1997) 2582-2588.
[23] J. Beens, M. Adahchour, R.J.J. Vreuls, K. van Altena, U.A.T. Brinkman, Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 919 (2001) 127-132.
[24] J. Harynuk, T. J. Gorecki, Comprehensive two-dimensional gas chromatography in stop-flow mode, J. Sep. Sci. 27 (2004) 432-441.
[25] J.V. Seeley, N.J. Micyus, S.V. Bandurski, S.K. Seeley, J.D. McCurry, Microfluidic deans switch for comprehensive two-dimensional gas chromatography, Anal. Chem. 79 (2007) 1840-1847.
[26] B.D. Quimby, J.D. McCurry, W.M. Norman, Capillary flow technology for gas chromatography: Reinvigorating a mature analytical discipline, LCGC The Peak (2007) 7-15.
[27] P.A. Bueno, J.V. Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1027 (2004) 3-10.
[28] J.V. Seeley, F. Kramp, C.J. Hicks, Comprehensive two-dimensional gas chromatography via differential flow modulation, Anal. Chem. 72 (2000) 4346-4352.
[29] A.E. Sinha, K.J. Johnson, B.J. Prazen, S.V. Lucas, C.G. Fraga, R.E. Synovec, Comprehensive two-dimensional gas chromatography of volatile and semi-volatile components using a diaphragm valve-based instrument, J. Chromatogr. A 983 (2003) 195-204.
[30] M. Poliak, M. Kochman, A. Amirav, Pulsed flow modulation comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1186 (2008) 189-195.

[31] Q. Gu, F. David, F. Lynen, K. Rumpel, G. Xu, P. Vos, P. Sandra, Analysis of bacterial fatty acids by flow modulated comprehensive two-dimensional gas chromatography with parallel flame ionization detector/mass spectrometry, J. Chromatogr. A 1217 (2010) 4448-4453.
[32] B.J. Prazen, K.J. Johnson, A.W. Weber, R.E. Synovec, Two-dimensional gas chromatography and trilinear partial least squares for the quantitative analysis of aromatic and naphthene content in naphtha, Anal. Chem. 73 (2001) 5677-5682.
[33] C.G. Fraga, B.J. Prazen, R.E. Synovec, Comprehensive two-dimensional gas chromatography and chemometrics for the high-speed quantitative analysis of aromatic isomers in a jet fuel using the standard addition method and an objective retention time alignment algorithm, Anal. Chem. 72 (2000) 4154-4162.
[34] J.V. Seeley, F.J. Kramp, K.S. Sharpe, S.K. Seeley, Characterization of gaseous mixtures of organic compounds with dual-secondary column comprehensive two-dimensional gas chromatography, J. Sep. Sci. 25 (2002) 53-59.
[35] VICI model DV-12-1116T, Valco Instruments Company Inc., Houston, TX, USA
[36] C.E. Freye, L. Mu, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1424 (2015) 127-133.
[37] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, J. Chromatogr. A 1255 (2012) 3-11.
[38] F. Westad, N.K. Afseth, R. Bro, Finding relevant spectral regions between spectroscopic techniques by use of cross model validation and partial least squares regression, Anal. Chim. Acta 595 (2007) 323-327.
[39] T. Rajalahti, O.M. Kvalheim, Multivariate data analysis in pharmaceutics: A tutorial review, Int. J. Pharm. 417 (2011) 280-290.
[40] A.A. Gowen, G. Downey, C. Esquerre, C.P. O'Donnell, Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients, J. Chemom. 25 (2011) 375-381.
[41] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1-17.
[42] J.S. Nadeau, R.B. Wilson, J.C. Hoggard, B.W. Wright, R.E. Synovec, Study of the interdependency of the data sampling ratio with retention time alignment and principal component analysis for gas chromatography, J. Chromatogr. A 1218 (2011) 9091-9101.
[43] B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based fisher ratio analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC–TOFMS) data using a null distribution approach, Anal. Chem. 87 (2015) 3812-3819.
[44] P.M. Harvey, R.A. Shellie, Data reduction in comprehensive two-dimensional gas chromatography for rapid and repeatable automated data analysis, Anal. Chem. 84 (2012) 6501-6507.
[45] T.J. Fortin, Assessment of variability in the thermophysical properties of rocket propellant RP-1, Energy Fuels 26 (2012) 4383-4394.
[46] G.D. Christian, Analytical Chemistry, 6th ed., John Wiley & Sons, Hoboken, NJ, 2004.
[47] H. van der Voet, Comparing the predictive accuracy of models using a simple randomization test, Chemom. Intell. Lab. Syst. 25 (1994) 313-323.

[48] ASTM Standard D2425-04, Standard Test Method for Hydrocarbon Types in Middle Distillates by Mass Spectrometry, ASTM International, West Conshohocken, PA, 2009.

Chapter 7. Gaining a Fundamental Understanding of Fuel Properties by Chemometric Analysis of Advanced Chemical Composition Measurement Data

This chapter was reproduced from C.E. Freye, M.C. Billingsley, R.E. Synovec, "Gaining a Fundamental Understanding of Fuel Properties by Chemometric Analysis of Advanced Chemical Composition Measurement Data" *prepared for submission in Energy & Fuels*

7.1 INTRODUCTION

Chemical composition of kerosene-based fuels is known to affect the fuel properties and performance [1-4]. Traditionally, multicomponent fuels, including middle distillate (kerosene) fractions and blends thereof, are analyzed for known impurities and hydrocarbon classes (e.g. sulfur, olefin, oxygenates, and aromatic compounds) [5]. Indeed it is a challenging endeavor to achieve precise control over the chemical composition of multicomponent distillate fuels due to variations in crude oil composition and place of origin, refinery and post-refinery conditions, or even the date and time the material was refined, treated, and formulated to meet the detail specification requirements. Gaining a better understanding between fuel composition and its relationship to physical properties (and ultimately performance) is vital for the design and analysis of energy conversion devices, including those used in aerospace propulsion. Indeed, much work has been performed aimed to establish correlations between the variation in fuel composition and fuel specifications [6-16]. In order to provide a sound connection between fuel composition and performance and physical properties, it has been shown to be beneficial to evaluate laboratory blends, where the analyst has relative control over chemical composition, as well as to assess the performance of "field" fuels [6-10,12-16]. However, as the number of fuels analyzed increases, a

reliable and straightforward method is needed in order to glean significant information while keeping analysis time short.

Gas chromatography (GC) is a traditional analytical technique that is amenable for the separation and analysis of volatile and semi-volatile mixtures [17,18]. When GC is coupled with mass spectrometry (MS), spectral information is gathered allowing for further selectivity and the ability to identify chemical compounds. GC – MS has been shown to be a powerful tool for the analysis of kerosene-based fuel [6-11]. Comprehensive two-dimensional (2D) gas chromatography coupled with time-of-flight mass spectrometry (GC × GC – TOFMS) can further improve upon the separation power of one dimensional GC and provide additional insight into complex mixtures of volatile compounds such as those present in kerosene-based fuels [13-15,19-21]. Traditionally in GC × GC, the first separation dimension uses a non-polar stationary phase while the second dimension uses a polar stationary phase. However, a reverse-column configuration with a polar first dimension and a non-polar second dimension has been shown to provide better selectivity for petroleum based samples [13-16,21]. A critical issue in the application of GC × GC – TOFMS is the sheer magnitude of data that is produced in the analysis of fuels, and the conversion of that data to useful information. This data analysis challenge is compounded by injection replicates for a large set of fuels. Yet, the challenge of obtaining meaningful information can be overcome through application of powerful chemometric software methods which can aid in the interpretation of such complex data sets [22,23].

Partial least squares (PLS) analysis is a chemometric method that associates the differences in measurable information for two different data sets. Details about PLS can be found elsewhere [24-27]. Briefly, PLS analysis mathematically relates, via linear algebra, two data matrices (X and Y block) through calculation of loadings referred to as latent variables (LVs). Using PLS, in the

study herein, models are constructed to account for the variance (ideally, the relevant chemical differences) in both the GC × GC – TOFMS data for a fuel sample set, i.e., the signal intensities (which constitute the X-block) and the respective measured property values, i.e., the ASTM data for the same fuel sample set (which constitute the Y-block). A key output of PLS modeling is a one-to-one correspondence between the measured ASTM values relative to the predicted ASTM values based upon the relationship to the chemical composition encoded in the GC × GC – TOFMS data, referred to as the PLS calibration. The other key output obtained from the PLS modeling is the linear regression vectors (LRVs) which indicates how the variables of the X-block (i.e. chromatographic and mass spectral information) relate to the Y-block (i.e. physical properties). PLS has been extensively used with GC × GC – TOFMS [15,28,29] and other GC × GC instrumental systems [16,30-34] to relate the chemical information to physical properties of interest. When constructing PLS models, it is important to have a large number of samples that are truly diverse in composition. For example, when the fuel sample set is derived from lab blends composed from a limited set of feed stocks, the fuel set is not truly diverse. Lack of true sample set diversity may lead to covariance between compounds (or classes of compounds) leading to misleading and inaccurate information provided by the LRVs in the PLS models, albeit with accurate PLS calibrations [15, 16]. In order to address this shortcoming, it is paramount to study a large, predominately *naturally* diverse fuel sample set.

Herein, we use an expanded fuel set of 74 kerosene-based fuels (i.e., jet and rocket fuels) with varying specifications (e.g. Jet A, JP-5, JP-8, RP-1, RP-2, etc.) in order to gain an accurately detailed insight into the relationship between selected fuel properties and chemical composition. Tables B.1 and B.2 (Supporting Information) summarize all the samples. Using PLS, the chemical information with the data obtained from GC × GC – TOFMS is correlated to the measured physical

properties of viscosity, heat of combustion, and hydrogen content as well as density at 15, 45 and 85 °C, respectively. An important goal is to expand upon the previous work to further gain insight and create robust PLS models with a large, diverse set of fuels [15,16]. Furthermore, we investigate the LRVs to gain a comprehensive, and accurate, understanding of the relationship of the physical properties to the chemical composition of the fuels.

7.2 EXPERIMENTAL

7.2.1 *ASTM Measured Values*

The physical properties for the 74 fuels (Tables B.1 and B.2, Supporting Information) were obtained in triplicate and the average value was used in the PLS modeling. The complete list of physical and compositional properties measured and methods used to obtain values are the following: viscosity (cSt) at 40 °C according to ASTM D445/D446; hydrogen content (mass%) according to ASTM D7171; heat of combustion (btu/lbm) according to ASTM D4809 using a Parr Instruments Model 6200 calorimeter; and density (g/ml) at 15, 45, and 85 °C according to ASTM D4052 using a Rudolph Research Analytical DDM2911 density meter. A summary of the property data is provided in Table B.2 (Supporting Information).

7.2.2 *GC × GC – TOFMS Analysis*

To investigate the chemical composition of the fuels, a GC × GC – TOFMS instrumental platform was used consisting of an Agilent 6890N GC (Agilent Technologies, Palo Alto, CA, USA), a thermal modulator (4D upgrade, LECO, St. Joseph, MI, USA), and a Pegasus III TOFMS (LECO, St. Joseph, MI, USA). Aliquots of the fuel samples were introduced to the GC × GC – TOFMS instrument via a 7683B auto-injector (Agilent Technologies, Palo Alto, CA, USA). The auto-injector was set to inject 1 µL of sample at a 200:1 split at an inlet temperature of 275 °C.

Prior to injection, HPLC grade acetone and hexane (Fisher Scientific) were used as solvent rinses. The primary GC × GC column (^{1}D) was a Rxi-17Sil MS: 29.5 m × 250 µm inner diameter (i.d.) × 0.25 µm film thickness, and the secondary GC × GC column (^{2}D) was a Rxi-1MS 1.5 m × 180 µm I.D. × 0.25 µm film thickness. Ultrahigh purity helium (Grade 5, 99.999%, Praxair, Seattle, WA, USA) was used as the carrier gas at a constant flow rate of 2.0 mL/min. The primary ^{1}D oven was held at 40 °C for 1.5 min before being temperature programmed at 5 °C/min to 200 °C where it was held for 1 min. The secondary ^{2}D oven was held at a +12 °C offset relative to the primary ^{1}D oven and the modulator block was held at a +30 °C offset to the primary ^{1}D oven. The modulation period was 3 s (separation run time of the ^{2}D column) with 0.75 s hot and cold pulses for each stage. The transfer line was set to 285 °C and the ion source was 225 °C. Mass channels, *m/z*, 35-334 at unit resolution were collected with an electron impact ionization voltage of 70 eV at 100 spectra/s after a 10 s acquisition delay. The 74 samples (Table B.1, Supporting Information) were analyzed with two GC × GC – TOFMS replicates for each sample. Samples 1-68 and 69-74 were analyzed in two different time periods separated by 6 months.

7.2.3 *Data Analysis*

The GC × GC – TOFMS chromatograms were imported into MATLAB 2015b (MathWorks, Natick, MA) using an in-house converting algorithm. The data were baseline corrected and binned by 4 data points (modulations) on the ^{1}D dimension and by 15 data points (150 ms) on the ^{2}D dimension, thereby reducing the sample matrix from (600 ^{1}D data points × 300 ^{2}D data points × 300 *m/z*) to (150 ^{1}D data points × 20 ^{2}D data points × 300 *m/z*). This binning procedure mitigated any minor run-to-run GC retention time misalignment while also reducing computation time [26]. Because Samples 69-74 were run 6 months later than Samples 1-68 there was minor ^{1}D retention time shifting. This shifting was readily corrected with a small linear

correction factor. The chromatographic data was then mean centered. After the preprocessing steps, the chromatographic data was forwarded into the PLS analysis which was performed using PLS Toolbox 8.02 (Eigenvector Research Inc., Wenatchee, WA, USA). The following physical properties were predicted: kinematic viscosity (cSt), and net heat of combustion (btu/lb), hydrogen content (mass%), and density (g/mL) at 15, 45, and 85 °C. Normally leave-one-out cross validation (LOOCV) is implemented with PLS [15,16,26,28,30], but with a large dataset, this would result in an excessively long computation time and potential overfitting of the data. Therefore, for construction of the PLS models, all underwent Venetian blinds cross validation for testing the overall predictive ability of the models and to determine the correct number of latent variables (LVs) [35,36]. For the Venetian blinds cross validation, a PLS calibration model was prepared with each group of N - 6 fuels, and the given physical property for the remaining 6 fuels was predicted, where N is the total number of fuels being analyzed. For example, if the PLS analysis was for 72 fuels (evenly divisible by 6), Samples 1, 14, 27, 40, 53, 66 would be left out of the calibration step (using all of the other fuels) but are then predicted, then Samples 2, 15, 28, 41, 54, 67 were left out of the calibration step and are then predicted, and so on in a round robin fashion. Since in this study of 74 fuels, N is not divisible by 6, then the first two groups of 6 excluded from the calibration step will have an additional fuel to make up for the difference (i.e. there will be one or more group of 7, not to exceed 5 groups). Hence in this study of 74 fuels, Samples 1, 14, 27, 40, 53, 66, 72 were left out of the PLS calibration step and then predicted, then Samples 2, 15, 28, 41, 54, 67, 73 were left out of the calibration step and then predicted, and so on. Linear regression vectors (LRVs), as defined in the PLS toolbox, were investigated to determine the relationship between the chemical composition and physical measurements. Analyte compounds that exhibit positive values in the LRV are correlated with increasing the specified

property, while negatively correlated analyte compounds exhibit negative values in the LRV and are correlated with decreasing the specified property [13-16, 24-27, 30]. The GC × GC – TOFMS chromatogram replicates were analyzed in separate PLS models and are overlaid on the same graph for ease of visualization. In order to determine the "goodness-of-fit" for the PLS models, the root mean square error of cross validation (RMSECV) was calculated as defined by,

$$RMSECV = \left[\frac{1}{N} * \sum(y_{i,cv} - y_{i,meas})^2\right]^{0.5} \qquad (7.18)$$

where N is the number of fuel samples, $y_{i,cv}$ is the cross validation predicted value of sample (i) and $y_{i,meas}$ is the measured value for the same sample. RMSECV can be normalized (NRMESCV) by dividing the RMESCV by the range of the $y_{,meas}$ values. In addition, principal component analysis (PCA) was performed on the three LRVs from the density measurements for the purpose of interrelating the three separate PLS models at 15, 45 and 85 °C.

7.3 RESULTS AND DISCUSSION

Using the GC × GC – TOFMS instrument in the reversed column configuration, an excellent 2D chromatographic separation of the fuels was achieved. Figure 7.1 shows the total ion current (TIC) chromatograms of four representative fuels. Figure 7.1A shows the chromatogram of Sample 1: YA2921HW10 which is a middle distillate cut fuel, mainly composed of alkanes and cycloalkanes. Figure 7.1B shows the chromatogram of Sample 4: WC0721HW01, a heavy distillate cut mainly composed of alkanes. Figure 7.1C shows the chromatogram of Sample 66: POSF 10264, a light distillate fuel primarily composed of alkanes and aromatics. Figure 7.1D shows the chromatogram of Sample 67: POSF 10289, a middle distillate fuel composed of alkanes, cycloalkanes, and aromatics. For each compound class (alkane, cycloalkane, and aromatic) it is possible to discern sub-classes (e.g. n-alkanes, isoalkanes, monocylcoalkanes, etc.). Figure B.1A

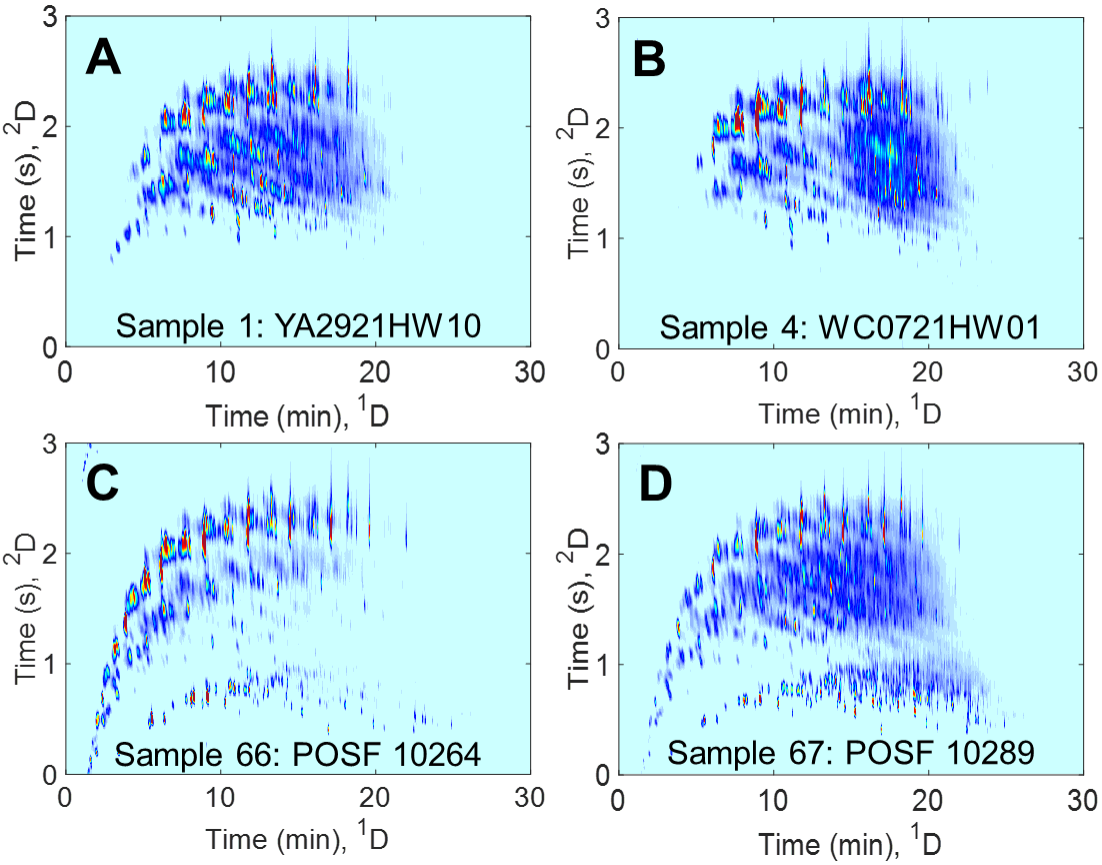


Figure 7.1 Four total ion current (TIC) GC × GC chromatograms representing the wide variety of fuels analyzed. (**A**) GC × GC chromatogram of Sample 1: YA2921HW10. (**B**) GC × GC chromatogram of Sample 4: WC0721HW01. (**C**) GC × GC chromatogram of Sample 66: POSF 10264. (**D**) GC × GC chromatogram of Sample 67: POSF 10289.

(Supporting Information) shows the approximate location in the 2D separation of the major compound classes: alkanes, cycloalkanes, and aromatics. It will be shown that each fuel can be thought of as providing its own chemical fingerprint comprised of a unique combination of these various compound classes and sub-classes. Note that if the data had not been collected using GC × GC (see Fig B.1B, Supporting Information), and a one-dimensional GC separation were employed, there would have been significant overlap between the compound classes making it difficult to distinguish them. Additionally, because each GC × GC chromatogram was collected

with a multivariate detector (TOFMS), it is possible to choose selective *m/z* that are representative of different chemical compound classes that exist in the fuels. To illustrate this point, Figure B.2A-F (Supporting Information) shows several analytical ion chromatograms (AIC) for Sample 67. This figure demonstrates the power of using selected *m/z* (as desired) coupled with the 2D separation advantage to isolate and in principle quantify the various hydrocarbon classes present in these fuels. However, in the current study, it was not necessary to identify and quantify every hydrocarbon class or every analyte present in the fuels in order to meet the goals of the PLS modeling.

For the sake of complete disclosure, the TIC GC × GC chromatograms of all 74 fuels is provided in Figure B.3 (Supporting Information). It should be noted that these 74 fuels have a varied range of chemical composition (unique fingerprint) with wide ranging distillate cut and various combinations of the major compounds classes (alkanes, cycloalkanes, and aromatics), with Figure 7.1 illustrating these differences. Furthermore, the fuels have various specifications (i.e. RP-1, RP-2, JP-8, Jet-A, etc.) along with a wide variety of fuel properties. There are nine fuels, Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72, that exhibit a significantly atypical chemical fingerprint compared to the majority of the fuels (remaining 65 fuels). The combination of the fuels provided in Figure 7.1 constitute the various compound classes exhibited by the "typical" fuel. Herein, we refer to any fuel that ultimately models well using PLS (Figure B.4, Supporting Information), yet is compositionally unique in its own right is referred to as a typical fuel. In contrast, Sample 48 (see Fig. B.3, Supporting Information), is only composed of 2 major peaks and several minor peaks. Another atypical fuel is Sample 16 (see Fig. B.3, Supporting Information) which at first glance may appear to be a typical fuel; however, the entire alkane region is not present and there are several excessively concentrated peaks. We shall see that the large chemical

variation present in these nine samples will make the PLS modeling of the fuels difficult, as is demonstrated next.

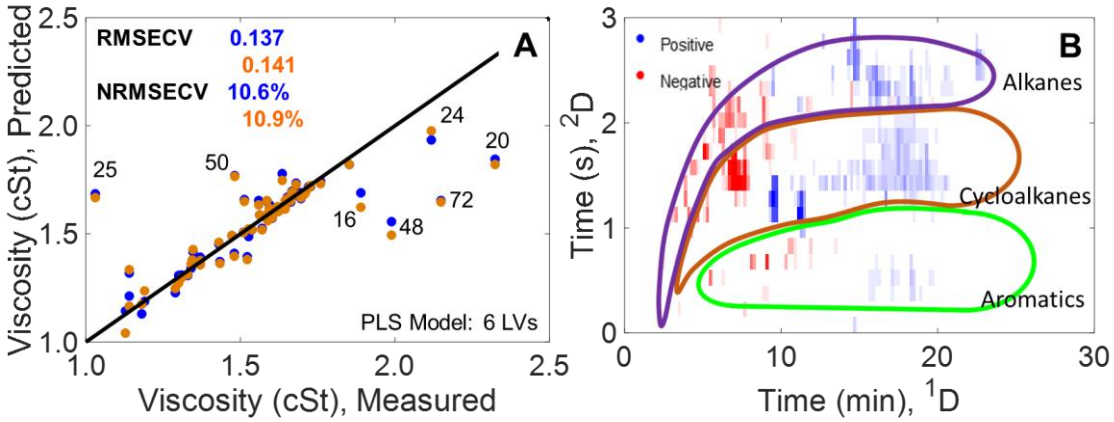


Figure 7.2 (A) PLS prediction of the viscosity using GC × GC – TOFMS chromatograms of all 74 fuel samples (Table B.1, Supporting Information). The black line represents the ideal agreement between the predicted and measured values. There are nine fuels that are incorrectly predicted (Figure B.4). (B) Linear regression vectors (LRVs) from the PLS prediction of the viscosity. Viscosity should be primarily based on distillate cut (i.e. boiling point). However, there appears to be 3 compounds eluting on the ^{1}D dimension at 8, 10, and 12 min that are dominating the model.

We now turn our attention to the PLS modeling of the physical properties. Using PLS modeling, the measured viscosity values (ASTM D445/D446) for all of the samples were predicted based upon the GC × GC – TOFMS chromatograms for each fuel. Figure 7.2A shows the calibration results from the modeled viscosity for all 74 fuel samples with both replicates of the GC × GC – TOFMS shown. The errors for the modeling are below a NRMSECV of 11%. However, there are several samples (Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72) which do not model correctly. Inspection of the LRV in Figure 7.2B indicate there are several chemical features,

three compound regions in particular, ${}^{1}t_{R} \approx 8$ min (large red feature), and 10 and 12 min (large blue features), which adversely influence the model the most. Investigating the Q residuals and Hoteling T^2 (see Figure B.4, Supporting Information) there are several fuels that adversely impact the PLS modeling due to large chromatographic differences relative to the typical fuel samples. This was mentioned earlier in the context of Figure 7.1 (and Figure B.3, Supporting Information) that some fuels stand out upon visual inspection as being "chemically" very different than a majority of the fuels. As previously mentioned, Sample 48 only has some minor peaks and two major peaks (chemical compounds) which are present in the sample at a grossly high concentration. Predicting the viscosity as well as the other physical properties will be adversely challenged because this sample is so chemically different from the rest of the fuels. Thus, fuels that exhibited a high Hoteling T^2, were excluded from the subsequent modeling of viscosity. There were a few fuels that had low Hoteling T^2 values, but moderately high Q residuals. Those fuels were kept in the subsequent modeling as they only led to a modest and acceptable increase in the modeling error.

In order to create a PLS model with low Q residuals and low Hoteling T^2, Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 were excluded, and viscosity modeling was re-evaluated with the remaining fuels (65 of the initial 74 fuels). The errors in the modeling are much lower with NRMSECV ~ 6% (instead of ~ 11%) in the calibration presented in Figure 7.3A with the average of both replicates shown, and with the fuels presented in Figure 7.1 highlighted. The PLS model with both replicates can be found in Figure B.5 (Supporting Information). The range of viscosity values over which the model accurately represents experimentally measured data is noteworthy. Additionally, the LRV in Figure 7.3B is more appropriate with no chemical compounds dominating the PLS modeling, and provides a sound chemical interpretation. Compounds with higher boiling points (longer retention times ${}^{1}t_{R}$ on the ^{1}D separation) are positively correlated with

increasing viscosity while lower boiling point compounds (shorter $^{1}t_{R}$ on the ^{1}D separation) are

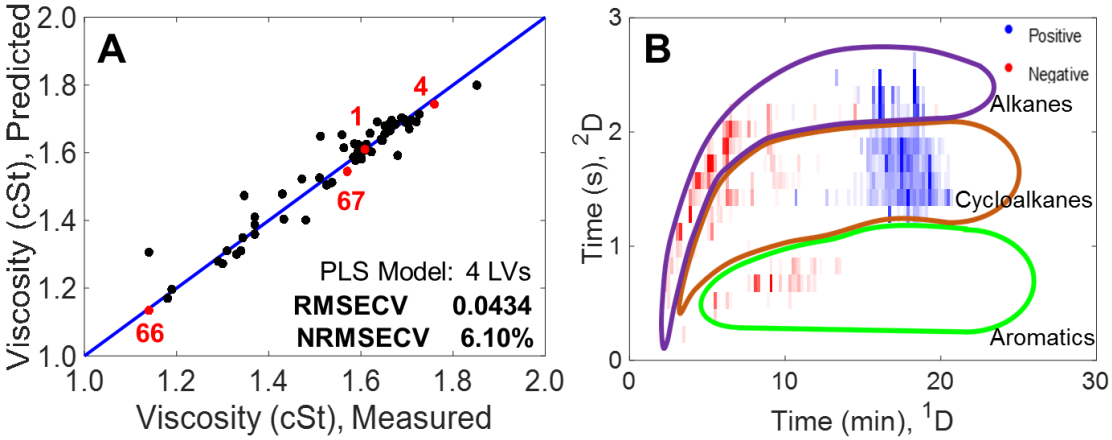


Figure 7.3 (A) PLS prediction of the viscosity using GC × GC – TOFMS chromatograms using 65 of the initial 74 fuels from Table B.1 (Supporting Information), with Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 objectively excluded based on statistical tests (Figure B.4). The blue line represents the ideal agreement between the predicted and measured values (slope = 1). Samples 1, 4, 66, and 67 have been highlighted red in order to discuss the relationship between their chemical composition, location on the calibration plot and linear regression vectors (LRV) (B) LRV from the PLS prediction of viscosity indicating that viscosity is a function of distillate cut (i.e. boiling point). Fuels that have a heavier distillate cut will have a higher viscosity.

correlated with lower viscosity. Comparing the chromatograms presented in Figure 7.1 to the LRVs, an analyst would expect Sample 4 to have a high viscosity, Samples 1 and 67 to have moderate viscosity, and Sample 66 to have a low viscosity. Sample 4 has more compounds that elute at longer $^{1}t_{R}$, Samples 1 and 67 have equal amounts of compounds that elute equally across the entire ^{1}D separation, and Sample 66 has a lack of compounds that elute at longer $^{1}t_{R}$. This is appropriately reflected in the calibration plot in Figure 7.3A.

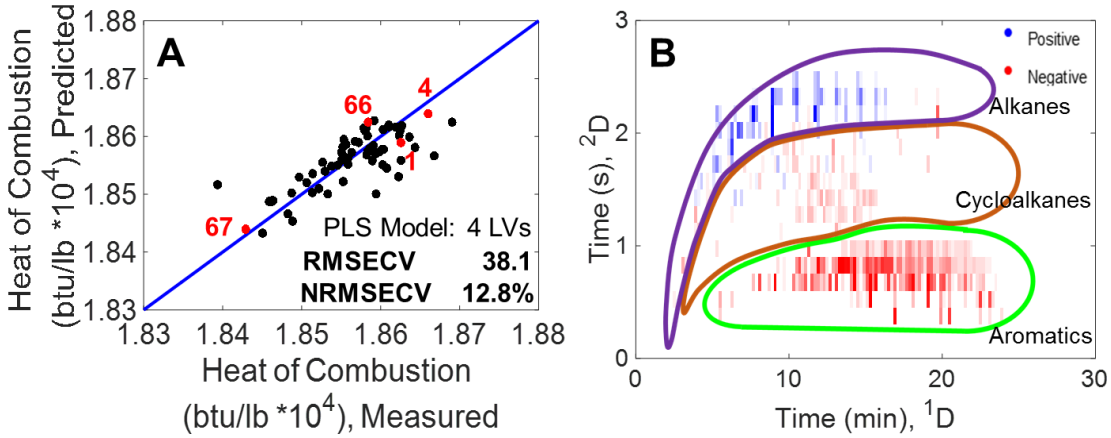


Figure 7.4 (A) PLS prediction of the heat of combustion using GC × GC – TOFMS chromatograms using 65 of the initial 74 fuels from Table B.1 (Supporting Information), with Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 objectively excluded based on statistical tests. The blue line represents the ideal agreement between the predicted and measured values. Samples 1,4, 66, and 67 have been highlight red in order to discuss the relationship between their chemical composition, location on the calibration plots and linear regression vectors (LRVs) (B) LRVs from the PLS prediction of heat of combustion indicating that heat of combustion is a dependent on the amount of alkanes, cycloalkanes, and aromatics present in the fuel. Fuels that have higher alkane content will have a higher heat of combustion, while fuels that have a large aromatic content will have a low heat of combustion. The cycloalkanes do not appear to effect the heat of combustion of a fuel relative to the contribution from alkanes and aromatics.

Before PLS prediction of the heat of combustion could be calculated, the heat of combustion values (see Table B.2, Supporting Information) were corrected for their hydrogen content (see Figure B.6, Supporting Information) which is a linear correction factor and accounts for the amount of heat lost to the water produced by the combustion [37,38]. Similar to the PLS prediction of viscosity, the PLS prediction of the heat of combustion is provided in Figure 7.4A,

and was carried out in a similar fashion to the viscosity modeling with the same outlier samples excluded: 16, 20, 24, 25, 46, 47, 48, 50, and 72. Samples mentioned above were excluded because they exhibited high hoteling T^2 values. The modeling of the heat of combustion has a slightly higher error than the viscosity and density (see below) with a NRMSECV of ~ 13%. However, the method precision (repeatability) and bias for ASTM D4809 are 0.096 and 0.089 MJ/kg (~ 40 btu/lb) respectively, and the model quality based on true values may therefore be better than indicated. Chemical interpretation of the LRV in Figure 7.4B indicates that the alkanes positively correlate with the heat of combustion while the aromatics negatively correlate with the heat of combustion. The cycloalkanes appear to not contribute to the heat of combustion based on the PLS modeling, at least not to the extent that the alkanes and aromatics do contribute. Comparing the chromatograms presented in Figure 7.1 to the LRVs, an analyst would expect Samples 1 and 4 to have a high net heat of combustion since no aromatics are present in the fuel, Sample 66 should have a moderately high net heat of combustion since there are a large amount of alkanes present but also a minor amount of aromatics, and Sample 67 should have a low net heat of combustion because there is a large amount of aromatics present. The calibration plot shown in Figure 7.4A matches this visual assessment. For comparison to Figure 7.4, Figure B.7 (Supporting Information) shows the PLS modeling of the heat of combustion with the inclusion of all fuels and the LRVs. In addition the PLS modeling of the heat of combustion excluding Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 is provided in which the replicates are plotted.

The PLS prediction of the hydrogen content was performed similarly to the viscosity and heat of combustion with the same samples being excluded (Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72) due to their high hotelling T^2 values. The calibration plot is provided in Figure 7.5A and has a NRMSECV error below 11%. Chemical interpretation of the LRVs (Figure 7.5B) indicates

that the alkanes are positively correlated while the aromatics are negatively correlated. In chemical terms, this matches as aromatics have pi bonds (double bonds) reducing the amount of saturation, and thus the amount of hydrogen present. It appears that the cycloalkanes do not appear to affect the hydrogen content significantly. Visually comparing the chromatograms present in Figure 7.1, Samples 1, 4, and 66 have a large amount of alkanes present in the fuel, Sample 66 has a moderate

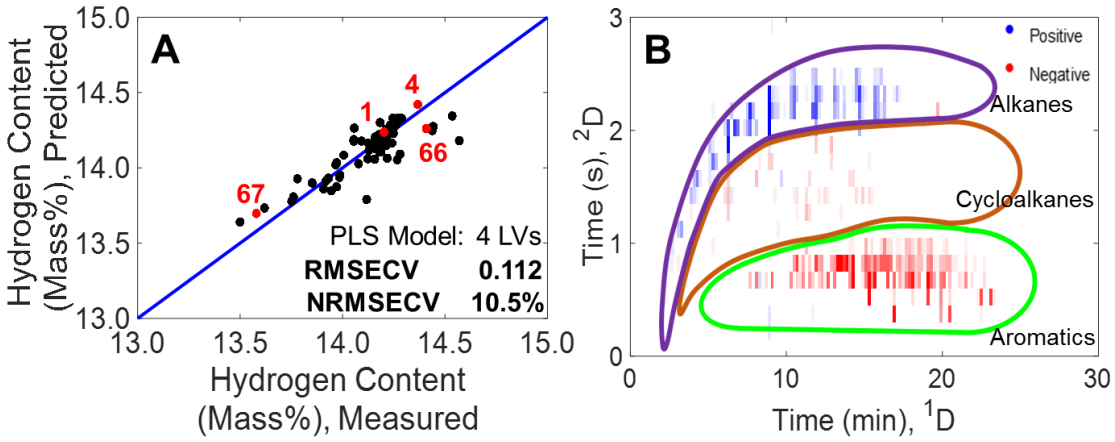


Figure 7.5 (A) PLS prediction of the hydrogen content using GC × GC – TOFMS chromatograms using 65 of the initial 74 fuels from Table B.1 (Supporting Information), with Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 objectively excluded based on statistical tests. The blue line represents the ideal agreement between the predicted and measured values. Samples 1,4, 66, and 67 have been highlight red in order to discuss the relationship between their chemical composition, location on the calibration plots and linear regression vectors (LRVs) (B) LRVs from the PLS prediction of hydrogen content indicating that hydrogen content is a dependent on the amount of alkanes and aromatics present in the fuel. Fuels that have higher alkane content will have a higher hydrogen content, while fuels that have a large aromatic content will have a low hydrogen content. It appears that the cycloalkanes do not seem to effect the hydrogen content of a fuel relative to the contribution from alkanes and aromatics.

amount of aromatics, and Sample 67 has a large amount of aromatics. Thus, Samples 1 and 4 should have a high hydrogen content, Sample 66 should have a moderate amount, and Sample 67 should have a low amount. However, Samples 1, 4, and 66 are all roughly predicted to have the same hydrogen content. One potential reason for Sample 66 having a relatively high hydrogen content could be the number of alkanes present in the sample are overriding the negative effects of having aromatics present. As with the other modeling, Figure B.8 (Supporting Information) shows the PLS analysis of the hydrogen content will the inclusion of all fuels and the resulting LRVs. Furthermore, the PLS modeling of the hydrogen content excluding Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 is shown with the replicates.

Finally, an overlay of the three PLS density models at 15 °C, 45 °C, and 85 °C is provided in Figure 7.6A with the average of both replicates indicated. The PLS model has an error of NRMSECV ~8%. Samples 66, 4, 1, and 67 (left to right in Fig 7.6A) have been highlighted for all three temperatures in order to discuss the relationship between the LRVs and chemical composition of the fuels. The LRVs for the PLS modeling of the density at 15 °C, 45 °C, and 85 °C are shown in Figures 7.6B-D, respectively. Interpretation of the LRVs indicate that alkanes are negatively correlated with density, cycloalkanes are moderately positively correlated with density, and aromatics are positively correlated with density. Sample 67 has a large amount of aromatics leading to a higher density, Sample 1 and 4 have no aromatics, approximately the same amount of cycloalkanes, but Sample 4 has been alkanes leading to a slightly lower density than Sample 1. Sample 66 has the lowest density as the same is mainly composed of alkanes with a minor amount of aromatics. It appears that density and heat of combustion are inversely correlated as compounds (or compound classes) that increase heat of combustion decrease the density of a fuel.

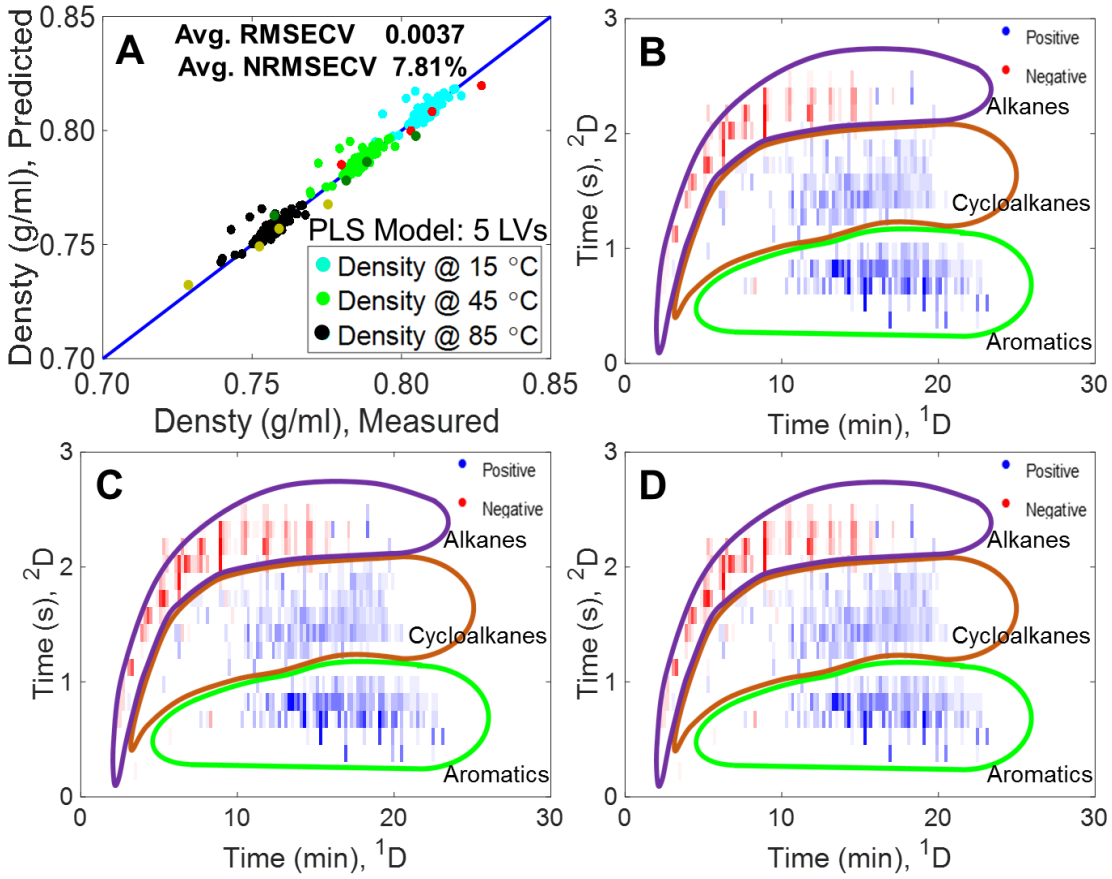


Figure 7.6 (A) Overlay of the three PLS prediction models of density at temperatures of 15 °C, 45°C, and 85 °C using GC × GC – TOFMS chromatograms with Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 excluded. The average of the two PLS modeling results for the two GC × GC – TOFMS replicates is shown. The blue line represents the ideal agreement between the predicted and measured values. Samples 66, 4, 1, and 67 (left to right) have been highlighted for each temperature so they can be discussed in the context of their chemical composition, location on the calibration plot and linear regression vectors (LRVs). (B) LRVs for PLS prediction of density at 15 °C. (C) LRVs for PLS prediction of density at 45 °C. (D) LRVs for PLS prediction of density at 85 °C.

The PLS evaluation of the density at 15, 45, and 85 °C including all the fuels can be seen in Figure B.9, B.10, and B.11 (Supporting Information), respectively. Additionally, PLS modeling of the density at 15, 45, and 85 °C excluding Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72 is shown with the replicates. The three LRVs in Figures 7.6B-D appear identical to each other as a function of temperature. In order to investigate this observation further, and to gain insight into the temperature dependence on density, the LRVs from each PLS modeling of the density (2 replicates × 3 different temperatures) were forwarded to PCA, with results presented in Figure 7.7. The LRVs separate on PC2 (PC1 not shown for brevity) in the scores plot in Figure 7.7A, forming a linear trend with respect to temperature. Over a sufficiently short range of temperature, density for a bulk sample such as kerosene based fuel is essentially linearly dependent on temperature. Investigating the loadings plot for the PCA in Figure 7.7B reflect what was originally observed in Figure 7.6, that the alkanes negatively influence density, while the cycloalkanes have a minor positive correlation with density, and aromatics are positively correlated. However, the loadings

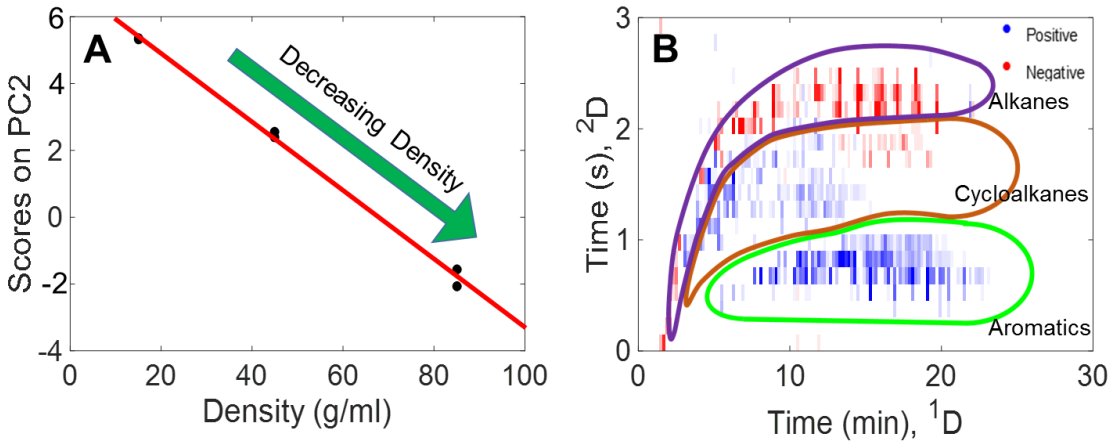


Figure 7.7 PCA modeling on the LRVs from the PLS prediction of density at 15, 45 and 85 °C, respectively. (A) Scores on PC2 form a linear function with respect to temperature. (B) Loadings plots from the PCA on the LRVs of the PLS prediction of density.

plot also provides insight into the relationship between density as a function of temperature. For each compound, density increases at a different rate (i.e. different slope in a function) which is caused by the intermolecular forces [39].

7.4 CONCLUSION

The overarching goal of this research is to better understand and enhance fuel performance through chemometric modeling, by relating fuel property data (ASTM measurements) to chemical composition data (GC × GC – TOFMS). The predictive PLS modeling approach presented herein may ultimately be used to educate fuel selection. Proof-of-principle of the PLS modeling approach was demonstrated on ASTM measurements of four physical properties: viscosity, heat of combustion, hydrogen content, and density. With an extremely large dataset composed of 74 diverse fuels, reliable PLS models were produced, that relate the chemical composition information provided by GC × GC – TOFMS measurements of each fuel to the physical properties of each fuel. Overall success of the approach relied heavily on having a compositionally diverse set of fuels with accurately measured physical properties. Additionally, it should be straight forward to transition this approach to either different fuels (jet fuel, diesel fuel, etc.) and/or different properties or even performance metrics such as thermal stability measurements. Future work should also focus on digging deeper into chemical compound identification that correlates to each physical property, following along the lines presented in Figure B.2 (Supporting Information).

7.5 SUPPORTING INFORMATION

Figures B.1 to B.11 and Table B.1 and B.2 can be found in Appendix B.

7.6 REFERENCES

[1] D.J. Cookson, B.E. Smith, Calculation of jet and diesel fuel properties using carbon-13 NMR spectroscopy, Energy Fuels. 4 (1990) 152–156. doi:10.1021/ef00020a004.

[2] G. Liu, L. Wang, H. Qu, H. Shen, X. Zhang, S. Zhang, Z. Mi, Artificial neural network approaches on composition–property relationships of jet fuels based on GC–MS, Fuel. 86 (2007) 2551–2559. doi:10.1016/j.fuel.2007.02.023.

[3] M.L. Huber, E.W. Lemmon, T.J. Bruno, Effect of RP-1 Compositional Variability on Thermophysical Properties, Energy Fuels. 23 (2009) 5550–5555. doi:10.1021/ef900597q.

[4] M.J. DeWitt, T. Edwards, L. Shafer, D. Brooks, R. Striebich, S.P. Bagley, M.J. Wornat, Effect of Aviation Fuel Type on Pyrolytic Reactivity and Deposition Propensity under Supercritical Conditions, Ind. Eng. Chem. Res. 50 (2011) 10434–10451. doi:10.1021/ie200257b.

[5] I.C. Lee, H.C. Ubanyionwu, Determination of sulfur contaminants in military jet fuels, Fuel. 87 (2008) 312–318. doi:10.1016/j.fuel.2007.05.010.

[6] M. Billingsley, T. Edwards, L. Shafer, T. Bruno, Extent and Impacts of Hydrocarbon Fuel Compositional Variability for Aerospace Propulsion Systems, in: 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, American Institute of Aeronautics and Astronautics, n.d. doi:10.2514/6.2010-6824.

[7] T.M. Lovestead, B.C. Windom, J.R. Riggs, C. Nickell, T.J. Bruno, Assessment of the Compositional Variability of RP-1 and RP-2 with the Advanced Distillation Curve Approach, Energy Fuels. 24 (2010) 5611–5623. doi:10.1021/ef100994w.

[8] R.V. Gough, T.J. Bruno, Composition-Explicit Distillation Curves of Alternative Turbine Fuels, Energy Fuels. 27 (2013) 294–302. doi:10.1021/ef3016848.

[9] P.Y. Hsieh, K.R. Abel, T.J. Bruno, Analysis of Marine Diesel Fuel with the Advanced Distillation Curve Method, Energy Fuels. 27 (2013) 804–810. doi:10.1021/ef3020525.

[10] J.L. Burger, T.J. Bruno, Application of the Advanced Distillation Curve Method to the Variability of Jet Fuels, Energy Fuels. 26 (2012) 3661–3671. doi:10.1021/ef3006178.

[11] N.J. Begue, J.A. Cramer, C. Von Bargen, K.M. Myers, K.J. Johnson, R.E. Morris, Automated Method for Determining Hydrocarbon Distributions in Mobility Fuels, Energy Fuels. 25 (2011) 1617–1623. doi:10.1021/ef101635a.

[12] T.J. Bruno, L.S. Ott, T.M. Lovestead, M.L. Huber, The composition-explicit distillation curve technique: Relating chemical analysis and physical properties of complex fluids, Journal of Chromatography A. 1217 (2010) 2703–2715. doi:10.1016/j.chroma.2009.11.030.

[13] M. Billingsley, N. Keim, R. Synovec, B. Hill-Lam, C. Wilhelm, Progress Toward Quality Assurance Standards for Advanced Hydrocarbon Fuels Based on Thermal Performance Testing and Chemometric Modeling, Air Force Research Laboratory (AFRL/RQRC) Edwards AFB United States, Air Force Research Laboratory (AFRL/RQRC) Edwards AFB United States, 2015. http://www.dtic.mil/docs/citations/AD1004638 (accessed February 7, 2018).

[14] M.C. Billingsley, N. Keim, B. Hill-Lam, R.E. Synovec, Hydrocarbon Fuel Thermal Performance Modeling based on Systematic Measurement and Comprehensive

Chromatographic Analysis, Air Force Research Laboratory (AFMC) AFRL/RQRC Edwards AFB United States, Air Force Research Laboratory (AFMC) AFRL/RQRC Edwards AFB United States, 2016. http://www.dtic.mil/docs/citations/AD1014607 (accessed February 7, 2018).

[15] B. Kehimkar, J.C. Hoggard, L.C. Marney, M.C. Billingsley, C.G. Fraga, T.J. Bruno, R.E. Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, Journal of Chromatography A. 1327 (2014) 132–140. doi:10.1016/j.chroma.2013.12.060.

[16] C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, Partial least squares analysis of rocket propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection, Talanta. 153 (2016) 203–210. doi:10.1016/j.talanta.2016.03.016.

[17] Z. Liu, J. B. Phillips, Comprehensive Two-Dimensional Gas Chromatography using an On-Column Thermal Modulator Interface, Journal of Chromatographic Science. 29 (1991) 227–231. doi:10.1093/chromsci/29.6.227.

[18] C.A. Bruckner, B.J. Prazen, R.E. Synovec, Comprehensive Two-Dimensional High-Speed Gas Chromatography with Chemometric Analysis, Anal. Chem. 70 (1998) 2796–2804. doi:10.1021/ac980164m.

[19] K.J. Johnson, R.E. Synovec, Pattern recognition of jet fuels: comprehensive GC×GC with ANOVA-based feature selection and principal component analysis, Chemometrics and Intelligent Laboratory Systems. 60 (2002) 225–237. doi:10.1016/S0169-7439(01)00198-8.

[20] C.G. Fraga, B.J. Prazen, R.E. Synovec, Comprehensive Two-Dimensional Gas Chromatography and Chemometrics for the High-Speed Quantitative Analysis of Aromatic Isomers in a Jet Fuel Using the Standard Addition Method and an Objective Retention Time Alignment Algorithm, Anal. Chem. 72 (2000) 4154–4162. doi:10.1021/ac000303b.

[21] B. Omais, M. Courtiade, N. Charon, D. Thiébaut, A. Quignard, M.-C. Hennion, Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate, Journal of Chromatography A. 1218 (2011) 3233–3240. doi:10.1016/j.chroma.2010.12.049.

[22] F. Westad, N.K. Afseth, R. Bro, Finding relevant spectral regions between spectroscopic techniques by use of cross model validation and partial least squares regression, Analytica Chimica Acta. 595 (2007) 323–327. doi:10.1016/j.aca.2007.02.015.

[23] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, Journal of Chromatography A. 1255 (2012) 3–11. doi:10.1016/j.chroma.2012.05.050.

[24] S.E. Prebihalo, K.L. Berrier, C.E. Freye, H.D. Bahaghighat, N.R. Moore, D.K. Pinkerton, R.E. Synovec, Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications, Anal. Chem. 90 (2018) 505–532. doi:10.1021/acs.analchem.7b04226.

[25] T. Rajalahti, O.M. Kvalheim, Multivariate data analysis in pharmaceutics: A tutorial review, International Journal of Pharmaceutics. 417 (2011) 280–290. doi:10.1016/j.ijpharm.2011.02.019.
[26] A.A. Gowen, G. Downey, C. Esquerre, C.P. O'Donnell, Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients, J. Chemometrics. 25 (2011) 375–381. doi:10.1002/cem.1349.
[27] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta. 185 (1986) 1–17. doi:10.1016/0003-2670(86)80028-9.
[28] S. Mostafapour, H. Parastar, N-way partial least squares with variable importance in projection combined to GC × GC-TOFMS as a reliable tool for toxicity identification of fresh and weathered crude oils, Anal Bioanal Chem. 407 (2015) 285–295. doi:10.1007/s00216-014-8076-1.
[29] J.R. Radović, K.V. Thomas, H. Parastar, S. Díez, R. Tauler, J.M. Bayona, Chemometrics-Assisted Effect-Directed Analysis of Crude and Refined Oil Using Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry, Environ. Sci. Technol. 48 (2014) 3074–3083. doi:10.1021/es404859m.
[30] V. Abrahamsson, N. Ristic, K. Franz, K. Van Geem, Comprehensive two-dimensional gas chromatography in combination with pixel-based analysis for fouling tendency prediction, Journal of Chromatography A. 1501 (2017) 89–98. doi:10.1016/j.chroma.2017.04.021.
[31] L.A.F. de Godoy, M.P. Pedroso, L.W. Hantao, R.J. Poppi, F. Augusto, Quantitative analysis by comprehensive two-dimensional gas chromatography using interval Multi-way Partial Least Squares calibration, Talanta. 83 (2011) 1302–1307. doi:10.1016/j.talanta.2010.08.015.
[32] H. Parastar, S. Mostafapour, G. Azimi, Quality assessment of gasoline using comprehensive two-dimensional gas chromatography combined with unfolded partial least squares: A reliable approach for the detection of gasoline adulteration, J. Sep. Science. 39 (2016) 367–374. doi:10.1002/jssc.201500720.
[33] B.J. Prazen, K.J. Johnson, A. Weber, R.E. Synovec, Two-Dimensional Gas Chromatography and Trilinear Partial Least Squares for the Quantitative Analysis of Aromatic and Naphthene Content in Naphtha, Anal. Chem. 73 (2001) 5677–5682. doi:10.1021/ac010637g.
[34] X. Shi, H. Li, Z. Song, X. Zhang, G. Liu, Quantitative composition-property relationship of aviation hydrocarbon fuel based on comprehensive two-dimensional gas chromatography with mass spectrometry and flame ionization detector, Fuel. 200 (2017) 395–406. doi:10.1016/j.fuel.2017.03.073.
[35] L. Ranzan, C. Ranzan, L.F. Trierweiler, J.O. Trierweiler, Classification of Diesel Fuel Using Two-Dimensional Fluorescence Spectroscopy, Energy Fuels. 31 (2017) 8942–8950. doi:10.1021/acs.energyfuels.7b00954.
[36] D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models. PLS-DA, Anal. Methods. 5 (2013) 3790–3798. doi:10.1039/C3AY40582F.
[37] J.J. McKetta, Encyclopedia of Chemical Processing and Design, CRC Press, 1990.

[38] A.C. Fieldner, W.A. Selvig, Use of the hydrogen-volatile-matter ratio in obtaining the net heating value of American coals, Govt. print. off., 1918.
[39] J.L. Trenzado, J.S. Matos, L. Segade, E. Carballo, Densities, Viscosities, and Related Properties of Some (Methyl Ester + Alkane) Binary Mixtures in the Temperature Range from 283.15 to 313.15 K, J. Chem. Eng. Data. 46 (2001) 974–983. doi:10.1021/je0100286.

Chapter 8. Recent Advances in the Development of Chemical Analysis Tools to Relate Compositional Variation to Thermal Integrity Data for RP-1, RP-2, and Related Fuels

This chapter was reproduced from C.E. Freye, B. Hill-Lam, N. Keim, M.C. Billingsley, and R.E. Synovec "Recent Advances in the Development of Chemical Analysis Tools to Relate Compositional Variation to Thermal Integrity Data for RP-1, RP-2, and Related Fuels" prepared for submission in *Energy & Fuels*

8.1 INTRODUCTION

The decomposition of hydrocarbons into carbonaceous deposits (also known as "coking") is a major problem found in the cracking of petroleum products [1-3], pyrolysis of biomass for biofuels [4-6], production of PVC [7,8], and regenerative cooling in rocket and jet engines [9-11]. The formation of carbonaceous deposits is a relatively well understood phenomenem, and it is generally accpeted that there are three basic mechanisms that govern the formation of carbon deposits [1,3,9,12-18]. These include the production of condensation (amorphous) carbon via formation of polycyclic aromatic hydrocarbons (PAH) in the hyrdocarbon stream which condense on the tubing, catalyst, and/or reactor wall surfaces, free radical growth reactions between the hyrdocarbon stream and the surface coke, and filamentous carbon via catalytic reactions with surface metals found in the tubing, catalyst, and/or reactor wall surfaces. Several studies using single component model compounds have shown that coking propensity as a function of chemical structure is n-alkanes < iso-alkanes < alkenes ≈ cycloalkanes < armomatics < PAH [3,16,19]. Furthermore, the the presence of heteroatoms such as those containing sulfur can increase coking rates [1,20].

The use of regenerative cooling in liquid-hydrocarbon-fueled hypersonic vehicles has been a staple for the aerospace field wherein the fuel flows through tubes or channels of a heat-exchanger in order to act as a heat-sink [12,21-24]. Robert Godard built the first regeneratively cooled engine in 1923, and subsquent design of the V-2 rocket, Rocketdyne F-1 and Rocketdyne J-2 Saturn V engines, Pratt & Whitney J58 engine for the SR-71, and many prior and successive engine designs use regenerative cooling. While regenerative cooling may be advantageous to reduce temperature of the combustion chamber, nozzle of the engine, and/or other sub-systems, a major shortcoming is the undesirable formation of carbonaceous deposits [9-12,25,26]. Investigation into the thermal stability of fuels has taken a three-fold approach: 1) fuel composition, 2) cooling surface substrates (i.e. copper, stainless steel, etc.), and 3) testing conditions (i.e. pressure and flow rates).

Using a single-tube flow reactor, two jet fuels, Jet A-1 and JP-8, and a Synthetic Paraffinic Kerosene (SPK) derived from the Fischer-Tropsch process were analyzed [27]. All three fuels were prone to pyrolysis and subsequent carbon deposition but the SPK was more reactive which was determined to be a result of its higher hydrogen content (%mass). Five common kerosene-based fuels: JP-7, JP-8, JP-8+100, JP-10, and RP-1 were evaluated using a restively heated tube test apparatus which simulates regeneratively cooled rocket engines [28]. Carbon deposition for each fuel was evaluated for both the stainless steel and copper tubing. All fuels deposited measureable amounts of carbon, however, the high sulfur content JP-8 and JP-8+100 testing with the copper tubing resulted in extremely high carbon deposition which forced the premature ending of the testing. This was mitigated by using stainless steel tubing. Comparing SS316 (Fe-rich alloy) and Inconel 600 (Ni-based superalloy) for the thermal stressing of Jet A fuel, deposits of carbon of similar composition were measured [29]. Coating the SS316 with an alumni-coating

significantly inhibited the deposition of carbon. This investigation was expanded by another group of who used eight different surface substrates [30]. Titanium-based coatings have also been shown to be effective at reducing carbon deposition [31]. It has been demonstrated that elevated pressures in regenerative cooling channels promotes carbon deposition wherein the amorphous carbon increased drastically which inhibited the formation of catalytic filamentous carbon [32].

Traditionally, multicomponent distillate fuels were analyzed for known impurities (e.g. sulfur, olefin, oxygenates, and aromatic compounds) and effects on fuel-wetted surfaces (e.g. copper corrosivity) [33], with additional testing in representative engine components and environments providing full assurance of system reliability. However, this approach can become expensive and time consuming as engine operating environments and the number of candidate fuel formulations increases, and skilled technicians may be required both for testing as well as analysis and interpretation of results. In order to evaluate fuel performance, an ideal approach is to implement an apparatus that exposes candidate fuels to perturbations in the most relevant environmental conditions occurring in rocket cooling system environments. By observing the response of this simulative environment through a number of measurable parameters, a candidate fuel's performance is assessed. Furthermore, the ability to correlate compositional information to these measurable differences yields the opportunity to more effectively select, control, and specify fuels for advanced applications.

It is important to note that the influence on certain types of compounds on fuel thermal performance is widely accepted. For example, in rocket grade kerosene, aromatic, olefinic, and sulfur-containing compounds are restricted to relatively low levels due to their detrimental influence on cooling performance and other engine behavior [34-38]. However, questions regarding existing specifications must be addressed before candidate fuels are adopted for use in

systems imposing high levels of thermal stress and surface reactivity. These include assessment of 1) validity of current test methods controlling deleterious compounds; 2) validity of current maximum allowable levels for contaminants and unsaturated species; 3) specific types of compounds in a given class or family, including their relative detriment to cooling performance; and 4) other compounds not governed by specification limits but suggested to influence fuel reactivity, decomposition, and deposit formation.

Gas chromatography (GC) is a traditional analytical platform to separate and analyze volatile and semi-volatile mixtures [39-41]. When coupled with mass spectrometry (MS), spectral information can be gathered allowing for further selectivity. Although GC – MS is a powerful technique, it cannot sufficiently separate complex mixtures such as kerosene-based fuels with only a single dimension of GC [33,39,42-45]. Comprehensive two-dimensional gas chromatography followed by time-of-flight mass spectrometry (GC × GC – TOFMS) can further improve upon the separation power of one dimensional GC and provide additional insight into complex mixtures of volatile compounds such as those present in kerosene-based fuels [45-52]. With GC × GC, two complementary separation dimensions are implemented to maximize the separation power, which allows the analyst to substantially improve upon the information gleaned from complex samples. Typically, the first separation dimension uses non-polar stationary phase while the second dimension uses a polar stationary phase. However, a reverse-column configuration is implemented herein, which provides better selectivity for samples, in which the first dimension column has a polar stationary phase and the second dimension column has a non-polar stationary phase [46,53,54]. Although multidimensional chromatography has recently experienced increased applicability for aerospace fuel analysis, the extent of its utility has heretofore been limited in most cases to qualitative comparisons between fuels and in some cases hydrocarbon type classification,

owing to the immensity of the datasets offered by GC × GC. With the amount of chemical information obtained from GC × GC – TOFMS, there are significant challenges in gleaning useful information from the data.

Partial least squares (PLS) analysis is well suited to relate complex chemical information obtained from GC × GC – TOFMS to other forms of chemical/physical measurements [45,46,52,55-58]. PLS yields two important outcomes, a linear correspondence of the chemical/physical properties to the GC × GC – TOFMS data. This can subsequently be used to predict modeled chemical/physical properties without having to directly measure these properties in new samples. The second outcome is the underlying relationship between the chemical composition of the samples and the predicted chemical/physical measurements. This is commonly denoted as the linear regression vectors (LRVs). In order to achieve reliable, robust PLS modeling, as well as determine the correct number of latent variables, cross validation is performed [45,46,52,54,56-58].

Herein for 36 fuels listed in Table 8.1, chemical information obtain via GC × GC - TOFMS is related to the physical properties obtained via CRAFTI analysis using PLS. The physical properties obtained from CRAFTI analysis and the subsequent LECO RC612 carbon deposition determinations were change in pressure as function of time and the carbon deposition profiles. A major goal is to determine if there are informative connections between the chemical information and the measured physical properties. Furthermore, we attempt to identify chemical classes or compounds that are predictive of the physical properties. Based on the information gained, improvements in the selection, control, and specification of future kerosene-based rocket propellants will be further enhanced. Herein, we aim to establish an experimental platform using advanced analytical techniques with chemometric data analysis to interrelate the chemical and

physical properties of kerosene-based rocket propellant, to begin to facilitate assessment of these questions.

8.2 EXPERIMENTAL

8.2.1 *CRAFTI and Carbon Deposition Analyses*

The thermal stability of the 36 fuels (Table 8.1) was probed using a instrument used to simulate regenerative cooling channel environments in liquid rocket engines. Figure C.1 (Supporting Information) shows pictures of the CRAFTI instrument. Fuel at a rate of $2.43*10^{-4}$ kg/min was flowed through a 30.5 cm copper tube which is termed the test article. The inner and outer dimensions of the test article are 0.91 and 3.18 mm, respectively. The inlet of the test article was held at 6.9 MPa. After the pressure at the inlet stabilized at 6.9 MPa, the test article was restively heated with an input power of 4500 Watts between 5.1 cm and 15.3 cm for 15 minutes. Pressure transducers were placed directly before and after the test article to measure the pressure. After the CRAFTI analysis, the test article was cut into 1.27 cm lengths. Each section was wiped with KimWipes and then immersed in an IPA bath. Each section was then evaluated using a LECO RC612 Carbon Determinator (LECO, St. Joseph, MI, USA). The furnace was held at 275 °C for 200 seconds, ramped to 400 °C at 100 °C/min and held for 120 seconds, ramped to 600 °C at 100 °C/min and held for 120s, and rampled to 900 °C at 100 °C/min and held for 60 seconds.

8.2.2 *GC × GC - TOFMS Analysis*

To investigate the chemical composition of the fuels a GC × GC – TOFMS instrumental platform was used consisting of an Agilent 6890N GC (Agilent Technologies, Palo Alto, CA, USA), a thermal modulator (4D upgrade, LECO, St. Joseph, MI, USA), and a Pegasus III TOFMS (LECO, St. Joseph, MI, USA). Aliquots of the fuel samples were introduced to the GC × GC –

TOFMS instrument via a 7683B auto-injector (Agilent Technologies, Palo Alto, CA, USA). The auto-injector was set to inject 1 µl of sample at 200:1 split at an inlet temperature of 275 °C. Prior to injection, HPLC grade acetone and hexane (Fisher Scientific) were used as solvent rinses. The primary GC × GC column (^{1}D) was a Rxi-17Sil MS: 29.5 m × 250 µm inner diameter (i.d.) × 0.25 µm film thickness, and the secondary GC × GC column (^{2}D) was a Rxi-1MS 1.5 m × 180 µm i.d. × 0.18 µm film thickness. Ultrahigh purity helium (Grade 5, 99.999%, Praxair, Seattle, WA, USA)

Sample Number	Sample Name	ΔP	Chemisorbed (Heated)	Chemisorbed (Exit)	Amorphous (Heated)	Amorphous (Exit)	Filamentous (Heated)	Filamentous (Exit)
1	YA2921HW10	61	11132	22324	18692	61411	3761	6234
2	BG1121GP04	139	221919	26964	37868	124977	8557	6593
3	GRC RP-1	68	10183	18488	20504	78314	5076	5038
4	WC0721HW01	62	9933	18169	15314	50304	3425	5155
5	LB073009-05	125	32812	18182	27719	195672	3948	8074
6	ZI1521HW10	65	7004	14275	10510	35407	2458	4080
7	CG0721HW10	60	6940	15767	11129	37951	2128	4190
8	LB073009-08	117	56238	22605	31728	185758	7307	7683
9	BB0821HW10	56	7279	16072	11896	53411	2462	4042
10	LB080409-05	57	7935	20909	13576	71105	2293	5159
11	ZI2621HW01	62	8356	17601	15792	42855	2684	5327
12	ZJ1321GP01	60	9606	20187	16469	41663	3060	5122
13	SA1421LS03	55	9207	19371	80050	114791	217066	5321
14	RG3021LS05	68	10302	16170	18106	81679	4875	5021
15	POSF 3327	80	20503	18060	13995	72899	2459	4297
16	POSF 4765	76	10248	12408	13008	50925	3246	3910
17	LB073009-02	142	12149	21712	33347	236008	3162	5742
18	VI2621LS01	57	7835	14196	10587	51332	2140	3700
19	DB131014	55	8583	19153	29843	66772	101423	5431
20	DC310925	65	25828	33114	39461	99501	634837	6045
21	DC310923	59	8466	21518	27106	116794	131354	5955
22	DB131013	52	8689	17831	13912	49201	3074	4985
23	DB131015	56	9979	22646	44914	107655	6705	4467
24	CL031236	57	7386	13319	11944	39685	2478	4612
25	EA130720	85	29865	25369	27150	179473	3560	7695
26	EB220705	58	10538	24095	16430	64247	6896	5030
27	CHC JP-5	65	11614	25561	31869	185521	4205	25103
28	LB080409-01	60	8811	15271	9228	51076	961	3948
29	LB073009-01	65	8777	16638	12523	70307	3032	4249
30	A0072256	68	9153	16228	12894	74876	2108	3611
31	CL11-3089	64	10232	16236	13512	77808	2307	3306
32	LB100413-40	61	8945	14684	11890	42479	2270	2898
33	LB073009-03	65	9487	17693	114567	49673	29524	3000
34	LB073009-10	63	9064	16388	13693	83661	2123	3931
35	ED060739	86	9770	16170	12238	67564	1613	2839
36	O11404	58	7200	15763	14381	49580	3998	3849

Table 8.1 Sample number and sample name are given. A Summary of the physical properties measured via CRAFTI and LECO's RC612 Carbon Determinator. The change in pressure is measured in psi and carbon measurements are in counts.

was used as the carrier gas at a constant flow rate of 2.0 ml/min. The primary ^{1}D oven was held at 40 °C for 1.5 min before being ramped at 5 °C/min to 200 °C where it was held for 1 min. The secondary oven was held at a +12 °C offset relative to the primary oven and the modulator block was held at a +30 °C offset to the primary oven. The modulation period was 3 s (separation run time of the ^{2}D column) with 0.75 s hot and cold pulses for each stage. The transfer line was set to 285 °C and the ion source was 225 °C. Mass channels, *m/z*, 35-334 at unit resolution were collected with an ionization voltage of 70 eV and detector voltage of 1850 at 100 spectra/s after a 10 s acquisition delay. The 36 samples (Table 1) were analyzed with two GC × GC – TOFMS replicates for each sample.

8.2.3 *Data Analyses*

The GC × GC – TOFMS chromatograms were imported into MATLAB 2015b (MathWorks, Natick, MA) using an in-house converting algorithm. The data was baseline corrected and binned by 3 data points (modulations) on ^{1}D and 20 data points (200 ms) on ^{2}D reducing the sample matrix to (200 ^{1}D data points × 15 ^{2}D data points × 250 *m/z*). This binning procedure mitigated any minor run-to-run GC retention time misalignment while also reducing computation time [46,52,54]. For the PLS modeling of the fuels, the 36 fuels were forwarded into PLS analysis which was performed using PLS Toolbox 8.02 (Eigenvector Research Inc., Wenatchee, WA, USA). The following physical properties were modeled: Δpressure, chemisorbed carbon in the heated zone, chemisorbed carbon in the exit zone, amorphous carbon in the heated zone, amorphous carbon in the exit zone, filamentous carbon in the heated zone, and filamentous carbon in the exit zone (Table 8.1). For construction of the PLS models, all underwent leave-one-out cross validation (LOOCV) for testing the overall predictive ability of the models and to determine the correct number of latent variables (LVs). The GC × GC – TOFMS chromatogram

replicates were analyzed in separate PLS models and are overlaid on the same graph for ease of visualization. In order to determine the "goodness-of-fit" for the PLS models, the root mean square

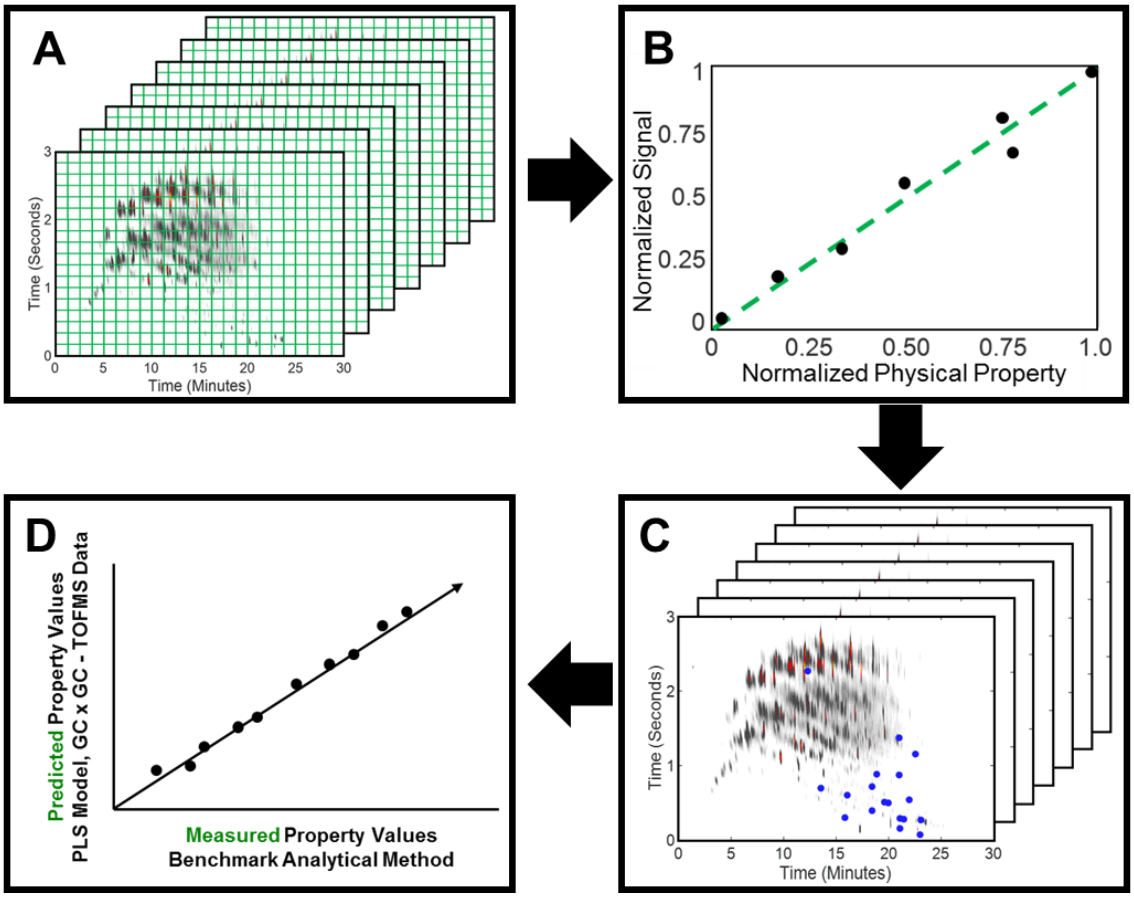


Figure 8.1 Illustration of the workflow for the PLS prediction of the physical properties. (A) For all fuel samples, GC × GC – TOFMS chromatographic data were binned in order to remove any chromatographic misalignment and reduce computation time. (B) Next, at each tile location, each mass channel was regressed against the physical property of interest using leave-one-out cross validation. (C) In order to be selected as a feature, at least 3 mass channels below a certain NRMSECV threshold were required. Mass channels that were above the threshold were not included in the final PLS modeling. (D) Mass channels at the tile locations that passed the thresholds were concatenated and regressed against the physical property using PLS.

error of cross validation (RMSECV) was calculated as defined by,

$$RMSECV = \left[\frac{1}{N} * \sum(y_{i,cv} - y_{i,meas})^2\right]^{0.5} \quad (8.19)$$

where N is the number of fuel samples, $y_{i,cv}$ is the cross validation predicted value of sample (i) and $y_{i,meas}$ is the measured value for the same sample. RMSECV can be normalized (NRMESCV) by dividing the RMESCV by the range of the $y_{,meas}$ values.

Next a novel feature selection technique was applied prior the calculation of the PLS models. It is been demonstrated previously that feature selection prior to PLS modeling can reduce errors and yield overall better models [59-62]; however, its application to GC × GC has been relatively limited [56]. Figure 8.1 shows the illustration of the workflow for the feature selection technique. As shown in Figure 8.1A, the chromatograms were binned by 3 data points on ^{1}D and 20 data points on ^{2}D reducing the sample matrix. At each tile, each *m/z* was individually regressed against the physical property of interest (Δpressure, chemisorbed carbon in the heated zone, chemisorbed carbon in the exit zone, amorphous carbon in the heated zone, amorphous carbon in the exit zone, filamentous carbon in the heated zone, and filamentous carbon in the exit zone) using LOOCV resulting in 750,000 regressions (200×15×250) as shown in Figure 8.1B. A normalized root mean square error of cross validation (NRMSECV) threshold was applied, and finally, to be deemed a relevant chemical feature, each tile identified had to contain at least 3 *m/z* below the NRMSECV threshold as shown in Figure 8.1C. The tiles that passed both of these requirements were then concatenated and used to holistically predict the physical property of interest using PLS as seen in Figure 8.1D.

8.3 RESULTS AND DISCUSSION

Using GC × GC – TOFMS in a reverse column configuration, an excellent 2D chromatographic separation of the compound classes: alkanes, cycloalkanes, and aromatics was achieved. Figure 8.2 shows the 2D separation of two representative fuels: Sample 1: YA2921HW10, and Sample 8: LB073008-08. The chromatograms of the other 34 fuels are excluded for brevity. The alkanes compounds are located from 2.0-3.0 s on ^{2}D, the cycloalkanes are located from 1.0-2.0 s on ^{2}D, and the aromatics are located form 0-1.0 s on ^{2}D and 5-30 min on ^{1}D. These regions were identified based on investigation of the mass spectra within each region and previous work [46,52,54]. Although it is possible to discern subclasses for the cycloalkanes and aromatic classes (i.e. monocyclics, dicyclics, tricyclics, mono-aromatics, and di-aromatics) and every hydrocarbon class (or each hydrocarbon) could be quantified and identified (utilizing retention indices and mass-to-charge ratios, *m/z*) in order to create compositional descriptions of the fuels based, we have elected to approach the chemical analysis of the fuels in a more targeted approach. The identification of the general elution times for the three classes (alkanes,

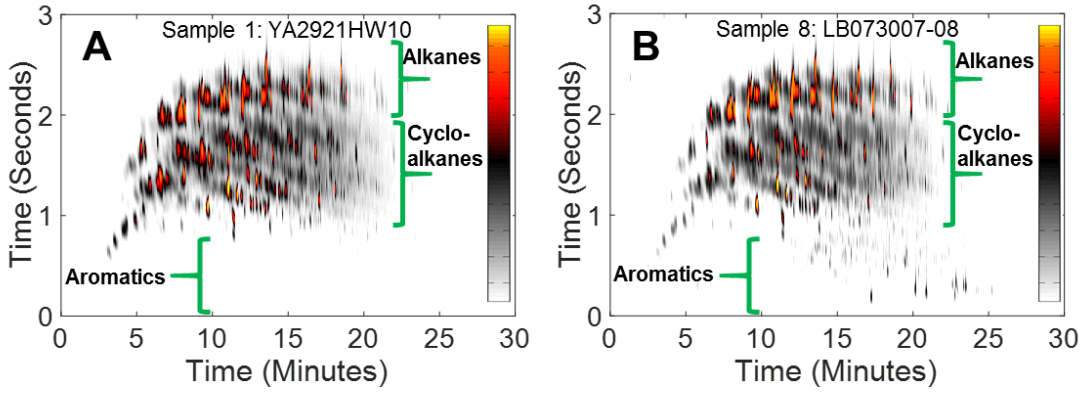


Figure 8.2 Shown are the two representative GC × GC - TOFMS chromatograms of kerosene-based rocket fuels (TIC – total ion current signal plotted). (A) Sample 1: YA2921HW10. (B) Sample 8: LB073007-08.

cycloalkanes, aromatics) in the 2D chromatographic space serves as a more instructive purpose of correlating general fuel chemical composition with differences in measured performance behavior. The key difference between Sample 1: YA2921HW10 (Fig. 8.2A) and Sample 8: LB073008-08 (Fig. 8.2B) is the presence of aromatic species. However, there are minor differences in the presence of alkane and cycloalkane chemical species. The 36 fuels exhibit a large composition range leading to their various physical properties and thermal stability.

The information obtained from the thermal integrity assessment (i.e. CRAFTI and LECO RC612 Carbon Determinator) is shown in Figure 8.3. Figure 8.3A shows the instrumental schematic for the test article of the CRAFTI instrument. The test article is divided into three major sections: the inlet, the heated zone, and exit zone. The change of pressure, defined as pressure at the inlet minus the pressure at the exit, as a function of time is shown in 8.3B. The pressure data from the multiple experimental tests for each of the 36 fuels was averaged together creating a single change of pressure vs. time vector for each of the fuels. There are 4 fuels that have a relatively large change in pressure as a function of time (2: BG1121GP04, 5: LB073009-05, 8: LB073009-08, and 17: LB073009-02), 4 fuels that have a very moderate change in pressure (15: POSF 3327, 16: POSF4765, 25: EA130720, and 35: ED060739), and the rest of the samples have relatively little change in pressure. However, it is important to note that all samples exhibit some change in pressure as a function of time. This is to be expected because as the test article temperature increases, the viscosity of the fuels changes. For the subsequent PLS modeling of the pressure, the measured value herein referred to as ΔP was defined as the change of pressure at 900 seconds minus the change of pressure at 0 seconds. These values for the 36 fuels are listed in Table 8.1. Figure 8.3C and 8.3D show the Carbon Deposition profiles (Carbon Chromatograms) for Sample 1: YA2921HW10 and Sample 8: LB073008-08, respectively.

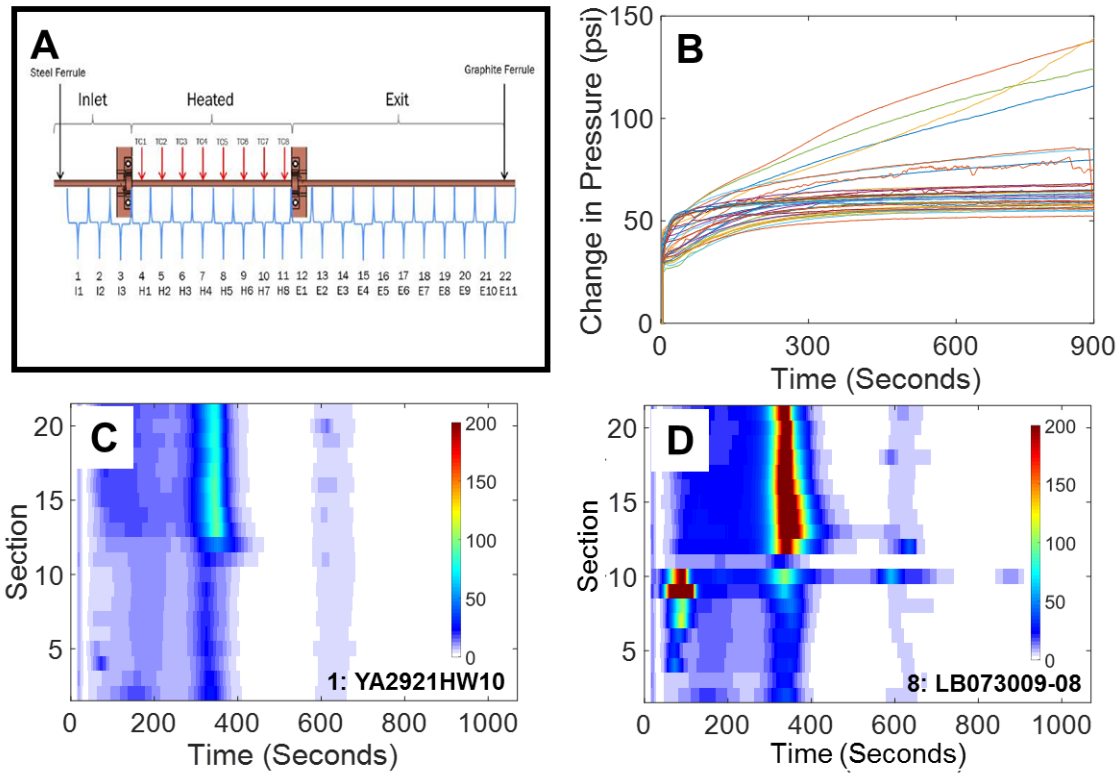


Figure 8.3 Thermal integrity information obtained from the CRAFTI analyses and subsequent Carbon Determinator analysis. (A) Schematic of the CRAFTI test article showing the three zones (inlet, heated, and exit) as well as the location of the thermocouples. (B) Pressure was measured at the beginning and end of the test article in order to investigate the backpressure generated by the deposition of the carbon. The change in pressure as defined as the difference of the pressure at the inlet minus the pressure at the outlet is shown as a function of time for all 36 fuels. (C) Carbon deposition for fuel YA2921HW10 is representative of fuels that exhibit low pressure changes. These fuels deposit a relatively small amount of chemisorbed and amorphous carbon. (D) Carbon deposition for fuel LB073007-08 is representative of fuels that perform poorly. These fuels deposit a relatively large amount of chemisorbed carbon in the heated zone (sections 4-11) as well as more amorphous carbon in the exit zone (sections 12-21).

As previously mentioned, sections 1 and 22 were excluded as these sections contained consistently high carbon content (~ 800°C oxidation temperature) due to the graphite ferrules used to seal the test article. Therefore, twenty sections per CRAFTI run were included in subsequent analyses. Four different types of carbon can be identified in the test article sections depending on the furnace temperature during oxidation analysis. Chemisorbed carbon is detected at ~ 190 °C, amorphous carbon is observed at ~ 385 °C, filamentous carbon is detected at ~600 °C, and crystalline graphite platelets are detected at ~ 800 °C. These furnace temperatures correspond to analysis times of approximately 100, 300, 625, and 900 seconds, respectively, in the Carbon Determinator two-dimensional data. For all of the fuels, there was limited deposition in sections 2 and 3 prior to the heated zone. Six fuels (2: BG1121GP04, 8: LB073008-08, 5: LB073009-05, 25: EA130720, 20: DC310925, and 15: POSF 3327) deposited large amounts of chemisorbed carbon in the heated zone (sections 4-11) and five fuels (17: LB073009-02, 5: LB073009-05, 8: LB073008-08, 27: CHC JP-5, and 25: EA130720) that deposited relatively large amounts of amorphous carbon in the exit zone (sections 12-20). All other fuels deposited small of amounts of chemisorbed carbon in the heated zone and amorphous carbon in the exit zone but at much lower counts. It is interesting to note that 4 fuels (20: DC310925, 13: SA1421L03, 21: DC310925, and 19: DB131014) deposited large amounts of filamentous carbon in the heated zone. Sample 13: SA1421L03 has been considered the "gold standard" for the industry due to fact that is has been through extra purification steps to remove any unwanted chemical species.

It is possible that the large deposition of filamentous carbon content is ironically due to this fuel being "too clean", with the copper surface of the test article serving as a catalyst for the filamentous carbon deposition growth analogous to producing carbon nanotubes [63]. More study is warranted to learn more regarding this possibility and its consequences for fuel composition

design and specification. However, this large deposition of filamentous carbon does not impact the change in pressure as function of time vectors. The carbon deposition values for the PLS modeling can be found in Table 8.1. In order to derive the carbon deposition values, the appropriate section number and time regions were summed to a single value. As previously mentioned, the heated zone and exit zone comprised sections 4-11 and 12-21, respectively. The chemisorbed carbon is located from 0 to 200 seconds, the amorphous carbon is located from 200 to 450 seconds, and the filamentous carbon is located from 450 to 700 seconds. With the large amount of information derived from the CRAFTI and Carbon Determinator analyses it was decided to originally focus on ΔP and the carbon values that correlated with ΔP. Figure C.2 (Supporting Information) shows the relationship between chemisorbed carbon in the heated zone (log scale and non-log scale) and ΔP and amorphous carbon in the exit zone and ΔP. Other figures of merit comparing ΔP and other forms of carbon are excluded for brevity but showed low correlation.

PLS was used to predict the ΔP based upon the entire chromatograms of the 36 fuels as shown in Figure C.3 (Supporting Information). The PLS model had a NRMSECV of ~17% but there were several fuels whose ΔP was predicted to be rather high when their measured ΔP was low. The PLS modeling for the logarithmically scaled chemisorbed carbon in the heated zone and amorphous carbon in the exit zone had similar errors (not shown for brevity). Because the PLS models with the entire dataset did not yield acceptable NRMSECVs and reliable models, a feature selection technique was implemented in order to reduce modeling errors. As described in the Experimental, the chromatograms were binned and then each *m/z* in each bin was regressed against the physical property using LOOCV resulting in 750,000 regressions. Figure 8.4A shows the 750,000 regressions for the feature selection for ΔP as slope as a function of NRMSECV. The slope was calculated by fitting a best fit line to the LOOCV regression. Unsurprisingly, as the

slope approaches 1, the NRMSECV approaches 0. Figure 8.4B shows the top 15,000 regressions. Figure C.4 (Supporting Information) shows *m/z* as a function of NRMSECV for the 750,000 regressions. It appears that there is no obvious relationship between *m/z* and NRMSECV and thus the feature selection thresholding (discussed below) was mainly based on the slope and NRMSECV values. Several NRMSECV thresholds were applied to the LOOCV regressions in order to better understand how to optimize the PLS modeling of the change in pressure. Table C.1 (Supporting Information) shows how the applied NRMSECV affects the number of features, number of tiles, and resulting error in the PLS modeling of the ΔP. Two NRMSECV thresholds of 14 and 17% were selected as this yielded a local minimum with a number of tiles great than 100. An analyst could opt to choose a lower NRMSECV threshold but this would result in a lower number of features leading to a less robust model.

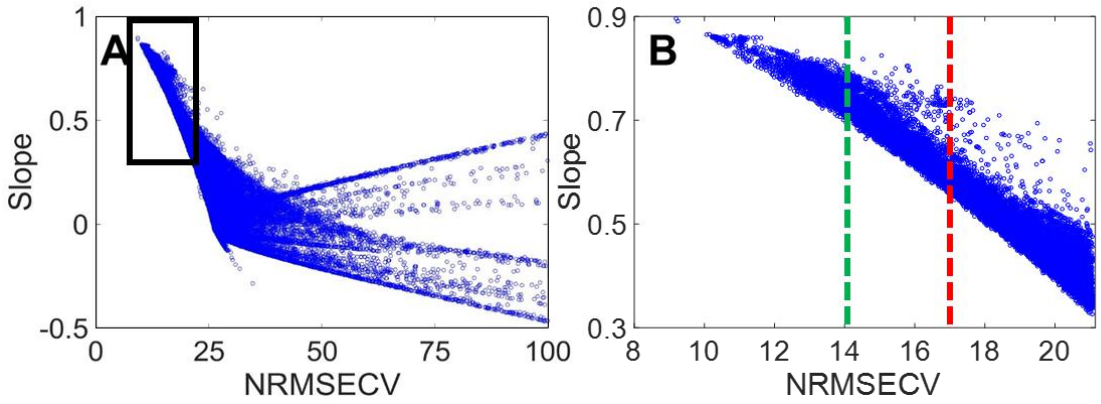


Figure 8.4 (A) The slope is plotted as a function of NRMSECV for the LOOCV regressions of the change in pressure. There are 750,000 mass spectra at the tile locations that were evaluated. Ideally, the evaluated mass spectra should have a slope close to 1 and a low NRMSECV. (B) An enhanced view of (A) is shown. Unsurprisingly, the lowest error in the regression has a slope closest to 1 which would indicate a good agreement between the predicted and measured change in pressure.

Figures 8.5A and 8.5B show the analytical ion chromatogram (AIC) of Sample 8 for the tiles identified by the LOOCV regressions for the change in pressure at a NRMSECV threshold of 14%. All of the tiles identified are located within the aromatic region of the chromatogram with a majority found within di-aromatic region (naphthalene derivatives). Figure 8.5C shows the

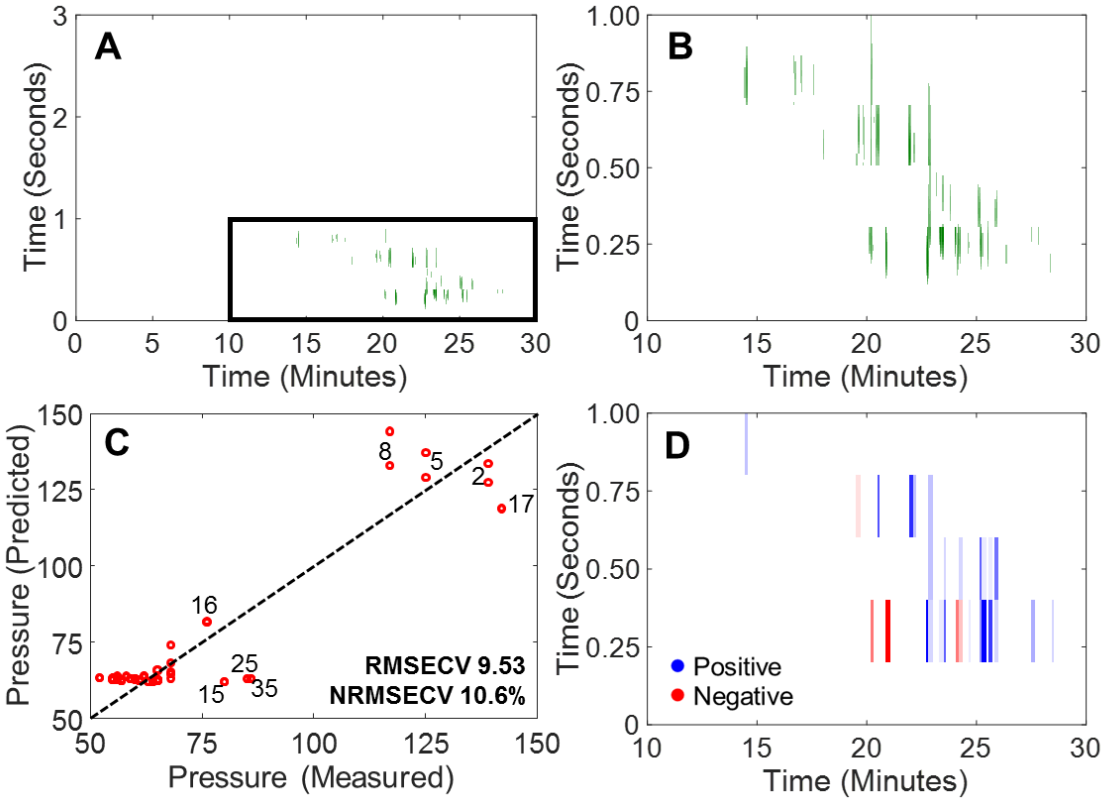


Figure 8.5 (A) The LOOCV regressions for the change in pressure from each tile per *m/z* which had a NRMSECV below 14% (RMSECV = 12.6 psi) and contained at least 3 *m/z* were used to create an analytical ion current chromatogram using Sample 8 to highlight the tiles identified. A total of 113 tiles were identified located within the aromatic compound and heteroatom containing compound portions of the chromatogram. (B) An enhanced view of A is shown. (C) PLS prediction of the change in pressure was performed using the tile locations and their respective *m/z* which were below the NRMSECV threshold of 14%. (D) Loadings of the PLS model indicating what tiles had a positive or negative correlation with ΔP.

subsequent PLS model based on those selected tiles and *m/z*. The resulting PLS NRMSECV error is 10.6% which is significantly lower than the non-feature selection PLS model which had a NRMSECV error of 16.9% and there are no significant outliers. The linear regression vectors

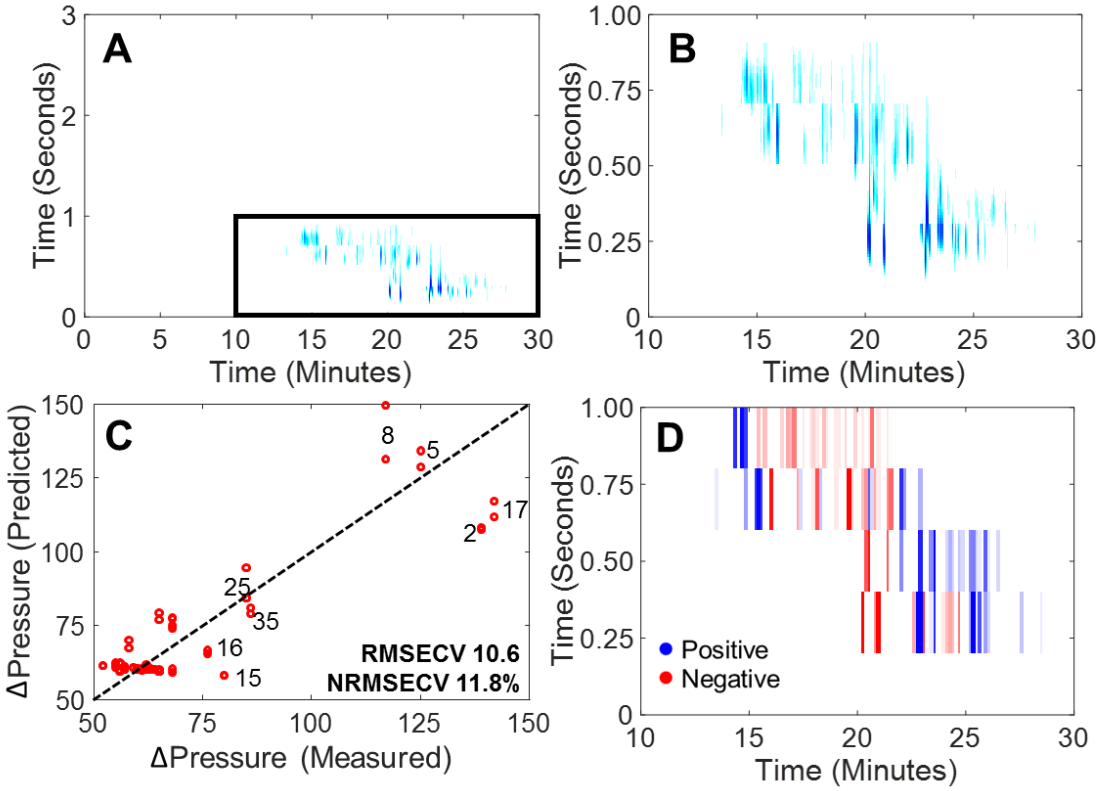


Figure 8.6 (A) The LOOCV regressions for the change in pressure from each tile per *m/z* which had a NRMSECV below 17% (RMSECV = 15.3 psi) and contained at least 3 *m/z* were used to create an analytical ion current chromatogram using Sample 8 to highlight the tiles identified. A total of 354 tiles were identified located within the aromatic compound and heteroatom containing compound portions of the chromatogram. (B) An enhanced view of A is shown. (C) PLS prediction of the change in pressure was performed using the tile locations and their respective *m/z* which were below the NRMSECV threshold of 17%. (D) Loadings of the PLS model indicating what tiles had a positive or negative correlation with ΔP.

(LRVs) of the PLS model of ΔP is shown in Figure 8.5D. The LRVs show which compounds are positively or negatively correlated with ΔP.

Figure 8.6A and 8.6B show the AIC of Sample 8 for the tiles identified by the LOOCV regressions for the change in pressure at a NRMSECV threshold of 17%. The identified tiles are similar to those when a 14% NRMSECV threshold was applied but there are significantly more tiles identified in between 0.5 and 1.0 seconds on ^{2}D. Figure 8.6C shows the PLS modeling of ΔP based upon the tiles that passed the 17% NRMSECV threshold. The PLS model has a similar error (11.8%) compared to the 14% threshold but closer inspection reveals that some of the fuels that exhibit high ΔP are farther away from the 1:1 line. Figure 8.6D shows the LRVs for the PLS model based on the 17% threshold. The LRVs are also similar to the 14% NRMSECV threshold LRVs but there are more tiles.

Using the tiled GC × GC – TOFMS data, the PLS prediction of the logarithmically scaled chemisorbed carbon in the heated zone and amorphous carbon in the exit zone was performed using the same approach as the ΔP. The figures showing the relationships between slope and *m/z* as a function of NRMSECV threshold have been excluded for brevity. Table C.2 and C.3 (Supporting Information) shows how the applied NRMSECV affects the number of features, number of tiles, and resulting error in the PLS modeling of the logarithmically scaled chemisorbed carbon in the heated zone and amorphous carbon in the exit zone, respectively. Figure 8.7A shows the tiles selected for the logarithmically scaled chemisorbed carbon in the heated zone for a NRMSECV threshold of 17%. A total of 86 tiles were identified and there is a striking similarity to the tiles identified by the ΔP at a threshold of 14% (Fig. 8.5A). The identified tiles were then used to holistically predict the logarithmically scaled chemisorbed carbon in the heated zone using

PLS as shown in Figure 8.7B. The PLS modeling has an NRMSECV error of 16.4%. Furthermore, the PLS model fails at predicting Sample 2 and underestimates Samples 15, 20, and 25.

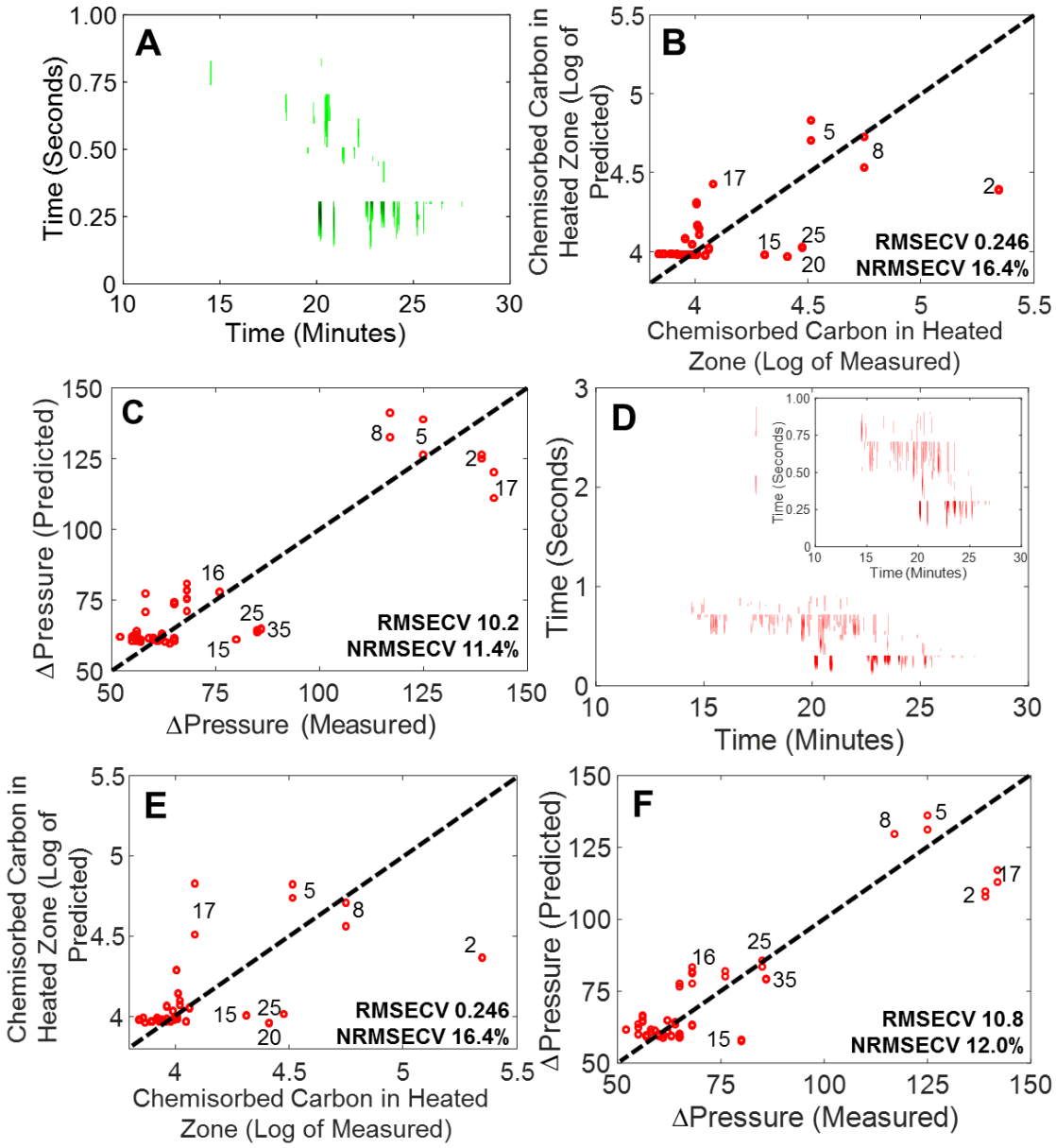


Figure 8.7 (A) The LOOCV regressions for the logarithmically scaled chemisorbed carbon in the heated zone (CCH) from each tile per *m/z* which had a NRMSECV below 17% (RMSECV = 0.2558) and contained at least 3 *m/z* were used to create an AIC using Sample 8 to highlight the tiles identified. A total of 120 tiles were identified which were predominantly within the aromatic portion of the chromatogram. (B) PLS prediction of the logarithmically scaled CCH. (C) PLS prediction of the ΔP based on the tiles selected in A. (D) The LOOCV regressions for the CCH from each tile per *m/z* which had a NRMSECV below 18% (RMSECV = 0.2709) and contained at least 3 *m/z* were used to create an AIC using Sample 8 to highlight the tiles identified. A total of 422 tiles were identified which were predominantly within the aromatic portion of the chromatogram. (E) PLS prediction of the logarithmically scaled CCH using the tiles identified in D. (F) PLS prediction of the ΔP based on the tiles selected in D.

Although these tiles did not perform a particularly well at predicting the logarithmically scaled chemisorbed carbon in the heated zone, the PLS prediction of the ΔP (Fig. 8.7C) is similar to using tiles specifically identified by the LOOCV regressions of the change in pressure. The PLS modeling error is low at 11.4% and is very similar to the PLS model of ΔP in Figure 8.5C. In addition, the LRVs (not shown for brevity) are extremely similar to the LRVs from the ΔP PLS model (Fig. 8.5D). Figure 8.7D shows the tiles identified when a NRMSECV threshold of 18% was used to select tiles which resulted in 422 tiles passing the threshold. The tiles identified are similar to those identified when a 17% threshold was applied to the ΔP. However, there are some tiles that are identified within the cycloalkane and alkane region that were not present in the ΔP.

Figure 8.7E shows the PLS modeling of the logarithmically scaled chemisorbed carbon in the heated zone based on the 18% NRMSECV threshold. The PLS model is nearly identical to 8.7B when a lower threshold was implemented. Although the chemisorbed carbon in the heated zone was logarithmically scaled, there may be other non-linear factors that are present that are causing the PLS modeling to fail. Using the tiles that passed the 18% NRMSECV threshold for the logarithmically scaled chemisorbed carbon in the heated zone, ΔP was PLS modeled. The PLS model has slightly higher errors compared to when a 17% NRMSECV threshold for the logarithmically scaled chemisorbed carbon in the heated zone, but is very similar to the PLS model of the ΔP when tiles were selected for ΔP when a threshold of 17% was used. However, the LRVs (not shown for brevity) are very different. Many of the compounds that were positively identified in Figure 8.6A are now negatively correlated with ΔP. Figure C.4A (Supporting Information) shows the AIC for Sample 8 for the tiles identified by the LOOCV regression of the amorphous carbon in the exit zone that were below a NRMESCV threshold of 17% and contained at least 3 *m/z* per tile. The location of these tiles is located within a different temporal location than the tiles identified by the LOOCV regressions of the logarithmically scaled chemisorbed carbon in the heated zone. Figure C.4B (Supporting Information) shows the PLS prediction of the amorphous carbon in the exit zone using the tiles identified. Relatively speaking, the PLS model does a reasonable job of predicting the amorphous carbon in the exit zone with the exception of Sample 25. Figure C.4C (Supporting Information) shows the PLS prediction of ΔP based on the tile identified by the feature selection for the amorphous carbon in the exit zone. This model significantly over predicts and under predicts Samples 27 and 2, respectively and has a higher error of 16.1%.

Tile Hit Number	^{2}D Tile	^{1}D Tile	Name	^2t_R (sec)	^1t_R (min)	Match Value	LRV Sign ΔP	# ΔP m/z	# CCH m/z
1	9	169	1,2,3-trimethylnaphthalene	0.25	25.4	923	+	83	93
			2-ethyl-5-methylbenzothiophene	0.22	25.4	870	+		
2	9	157	3,5-dimethylbenzothiophene	0.18	23.6	825	+	74	88
			1-ethylnaphthalene	0.32	23.6	869	+		
3	9	171	2-methyl-1,1'-biphenyl	0.27	25.7	949	+	63	71
4	9	152	Biphenyl	0.11	22.8	881	-	59	23
5	11	147	1-cyclohexyl-3-methylbenzene	0.58	22.1	908	+	56	10
6	9	162	2-isopropylnaphthalene	0.30	24.3	941	-	51	50
			1,5-dimethylnapthalene	0.23	24.3	954	+		
7	10	168	2-t-butylnapthalene	0.47	25.2	773	+	51	0
			4-methyl-1,1'-biphenyl	0.45	25.2	805	+		
8	9	161	1,3-dimethylnaphthalene	0.31	24.2	953	+	50	49
9	9	184	1,4-dimethylnapthalene	0.17	24.5	681	+	48	57
10	9	168	4-methyl-1,1'-biphenyl	0.26	25.2	958	-	46	67
11	9	156	3,5-dimethylbenzothiophene	0.20	23.4	872	+	42	75
12	10	173	1,7-dimethylnapthalene	0.20	24.3	956	+	40	0
13	9	186	1,4,5-trimethylnapthalene	0.22	27.9	862	+	38	42
14	9	140	2-methylnapthalene	0.14	21.0	956	+	33	33
15	9	176	2-methyl-diphenylmethane	0.20	26.4	917	+	32	39
			1,4,6-trimethylnapthalene	0.24	26.4	926	+		
16	11	153	6,7-dimethyltetralin	0.51	23.0	730	+	32	4
17	11	137	2,4-dimethylcyclopentylbenzene	0.56	20.6	774	+	27	36
18	9	175	2-isopropylnaphthalene	0.28	26.3	827	+	26	3
19	10	159	2-isopropylnaphthalene	0.37	23.9	900	-	22	7
20	10	171	1-butylnapthalene	0.37	25.7	760	+	21	0
21	9	190	1-Isopropenylnaphthalene	0.21	28.5	723	+	20	0
22	11	157	2,6-dimethylnaphthalene	0.51	23.6	768	-	20	0
23	12	153	4,5-di-epi-Aristolochene	0.85	23.0	749	+	19	3
24	12	135	Megastigma-4,6(E),8(Z)-triene	0.79	20.3	789	+	17	6
25	9	177	Dibenzofuran	0.13	26.6	929	+	15	53
26	12	97	1-Isobutyl-2,5-dimethylbenzene	0.74	14.6	921	+	15	7
27	9	136	2,7-dimethyltetralin	0.43	20.4	855	+	14	0
28	11	133	1,1,6-trimethyltetralin	0.55	20.0	846	+	14	13
29	9	153	2-ethylnaphthalene	0.20	23.0	944	+	13	65
30	9	170	1,4,6-trimethylnapthalene	0.31	25.5	846	-	12	30

Table 8.2 List of compounds identified by the LOOCV regressions of the change in pressure, and the logarithmically scaled chemisorbed carbon in the heated zone. The top 30 tiles are shown. A comprehensive list for all tiles and for all the forms of carbon are shown in Table C.3 (Supporting Information).

Based on the temporal location of the tiles for the chemisorbed carbon in the heated zone and the tiles for amorphous carbon in the exit zone compared to the tiles for the ΔP and the subsequent PLS modeling of the ΔP (Fig 8.5C, Fig 8.7C, and Fig C.5, Supporting Information), it is concluded that the deposition of chemisorbed carbon in the heated zone is the most responsible for the change in pressure (Fig 8.3B). In order to provide a metric to the correlation between the tiles identified by the ΔP and the various forms of carbon and locations, the number of common tiles identified in the LOOCV regressions were compared. Table 8.2 provides a list of the top 30 tiles identified by the LOOCV regressions of ΔP ordered in rank of number of *m/z*. The number of *m/z* identified by the chemisorbed carbon in the heated zone LOOCV regressions is included. In addition, tentative identification of the compounds contained within the tiles identified by the LOOCV regressions of ΔP are included. Many of the tiles contained two chemical species. Most of the compounds identified are naphthalene derivatives but it is important to note that there were several sulfur and oxygen containing compounds were found to be highly correlated to ΔP and the chemisorbed carbon in the heated zone. A comprehensive list for all the different forms of carbon and locations is shown in Table C.3. For sake of brevity identification of analytes in all the tiles are not included. Of the 113 tiles identified by the ΔP LOOCV regressions, the logarithmically scaled chemisorbed carbon in the heated zone had 74 tiles in common or 66%. The chemisorbed carbon in the exit zone had 6%, the amorphous carbon in the heated zone had 8%, the amorphous carbon in the exit zone had 15%, the filamentous carbon in the heated zone had 1% and the filamentous carbon in the exit zone had 5%. While there are a large number of tiles in common between the ΔP and the logarithmically scaled chemisorbed carbon in the heated zone and a minor amount of tiles between the ΔP and the amorphous carbon in the exit zone, there are a large amount of tiles (compounds) that are not shared between the carbon locations and the ΔP. Potential reasons

for this phenomenon are tiles (compounds) may be logarithmically (or another form of scaling) related to carbon but linearly related to carbon. Further investigation into this is needed.

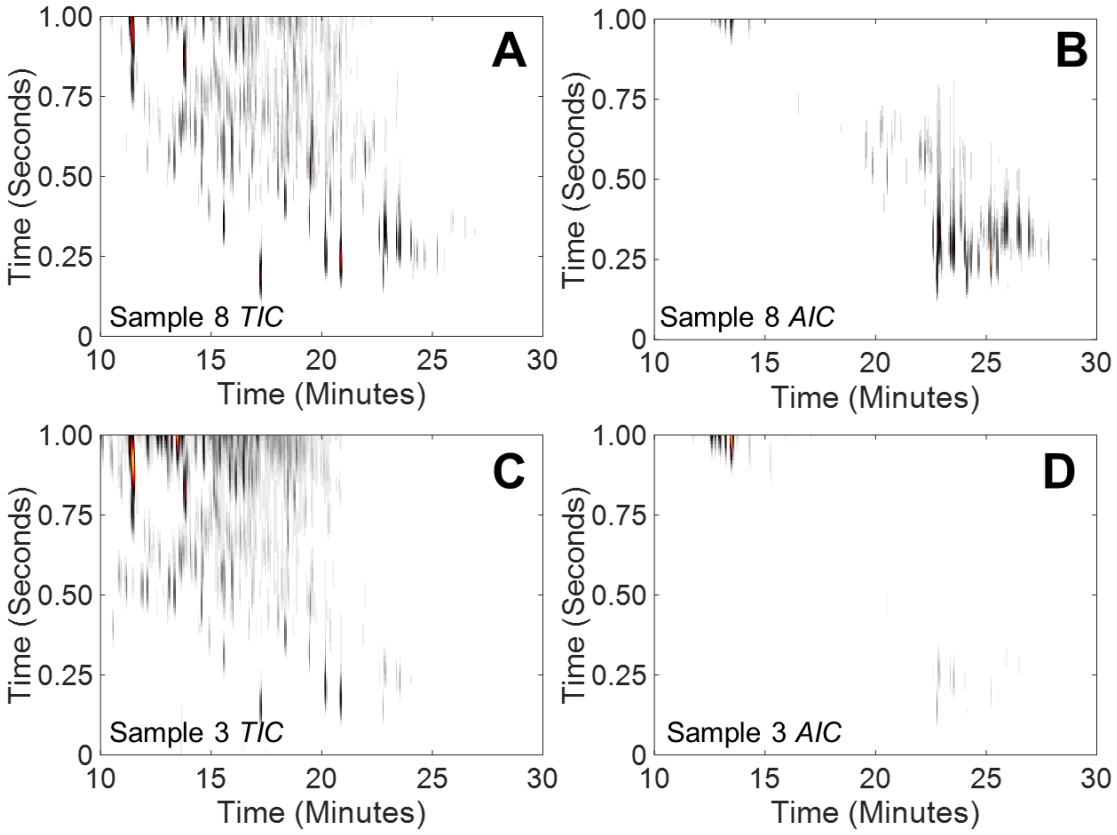


Figure 8.8 (A) TIC for Sample 8 showing a zoomed in area. (B) AIC for Sample 8 highlighting compounds that were identified by the LOOCV regressions of ΔP. (C) TIC for Sample 3 showing a zoomed in area. (D) AIC for Sample 3 highlighting compounds that were identified by the LOOCV regressions of ΔP.

In order to highlight some of the compounds identified by the feature selection process, Figure 8.8 has been prepared. The TIC GC × GC - TOFMS of Sample 8 is shown in Figure 8.8A and the AIC of Sample 8 prepared using *m/z* 152, 157, 167, and 170, which are highly selective for the majority of the compounds identified by the feature selection is shown in Figure 8.8B. The TIC GC × GC - TOFMS of Sample 3 is shown in Figure 8.8C and the AIC of Sample 3 prepared

using *m/z* 152, 157, 167, and 170 which is shown in Figure 8.8D. Using the entire chromatogram of Sample 3 results in overprediction of the ΔP (Fig C.3, Supporting Information). Visually comparing the TIC chromatograms of Sample 3 and Sample 8, they appear to be very chemically similar, but closer inspection of the AIC in Figure 8.8B and 8.8D indicates that Sample 3 is missing almost all of the compounds that were identified by the feature selection.

8.4 CONCLUSION

Herein we have demonstrated the establishment of a fuel analysis platform using advanced analytical techniques and chemometrics to obtain useful information regarding the chemical and physical properties of kerosene-based rocket propellant, as these properties relate to thermal integrity. GC × GC – TOFMS data, CRAFTI pressure data, and Carbon Determinator data all provide insight into fuel properties. A new chemometric technique based on PLS analysis was developed that readily identifies locations in the chromatograms which correlate with measured CRAFTI and Carbon Determinator data. Using PLS, the chemical data was used to predict useful physical property information. For the ΔP, the logarithmically scaled chemisorbed carbon in the heated zone, based upon PLS modeling with the GC × GC – TOFMS data, important chemical compounds have been identified. The findings of this research suggest that the chemisorbed carbon in the heated zone correlates strongly with the ΔP. Further investigation is warranted into the other measured physical properties: chemisorbed carbon in the exit zone, amorphous carbon in the heated zone, and filamentous carbon in the heated zone and exit zone. There were several fuels which deposited these types of carbon but had a low measured change in pressure. This prompts a further investigation into the chemical composition of the fuels and how chemical composition impacts the physical response in the CRAFTI testing. Using the information gleaned from the CRAFTI testing coupled with GC × GC – TOFMS analysis, chemical composition of kerosene

rocket fuel can further be tailored to achieve optimal physical properties. Additional studies are warranted to provide a deeper insight into how the chemical properties affect the physical properties of these fuels.

8.5 SUPPORTING INFORMATION

Figures C.1 to C.5 and Table C.1 and C.4 can be found in Appendix C.

8.6 REFERENCES

[1] M.-F.S.G. Reyniers, G.F. Froment, Influence of Metal Surface and Sulfur Addition on Coke Deposition in the Thermal Cracking of Hydrocarbons, Ind. Eng. Chem. Res. 34 (1995) 773–785. doi:10.1021/ie00042a009.

[2] G.C. Reyniers, G.F. Froment, F.-D. Kopinke, G. Zimmermann, Coke Formation in the Thermal Cracking of Hydrocarbons. 4. Modeling of Coke Formation in Naphtha Cracking, Ind. Eng. Chem. Res. 33 (1994) 2584–2590. doi:10.1021/ie00035a009.

[3] F.D. Kopinke, G. Zimmermann, G.C. Reyniers, G.F. Froment, Relative rates of coke formation from hydrocarbons in steam cracking of naphtha. 3. Aromatic hydrocarbons, Ind. Eng. Chem. Res. 32 (1993) 2620–2625. doi:10.1021/ie00023a027.

[4] R. French, S. Czernik, Catalytic pyrolysis of biomass for biofuels production, Fuel Processing Technology. 91 (2010) 25–32. doi:10.1016/j.fuproc.2009.08.011.

[5] Y. Zhang, S. Kajitani, M. Ashizawa, Y. Oki, Tar destruction and coke formation during rapid pyrolysis and gasification of biomass in a drop-tube furnace, Fuel. 89 (2010) 302–309. doi:10.1016/j.fuel.2009.08.045.

[6] S. Du, J. A. Valla, G. M. Bollas, Characteristics and origin of char and coke from fast and slow, catalytic and thermal pyrolysis of biomass and relevant model compounds, Green Chemistry. 15 (2013) 3214–3229. doi:10.1039/C3GC41581C.

[7] K. Zhou, J. Jia, X. Li, X. Pang, C. Li, J. Zhou, G. Luo, F. Wei, Continuous vinyl chloride monomer production by acetylene hydrochlorination on Hg-free bismuth catalyst: From lab-scale catalyst characterization, catalytic evaluation to a pilot-scale trial by circulating regeneration in coupled fluidized beds, Fuel Processing Technology. 108 (2013) 12–18. doi:10.1016/j.fuproc.2012.03.018.

[8] A.G. Borsa, A.M. Herring, J.T. McKinnon, R.L. McCormick, G.H. Ko, Coke and Byproduct Formation during 1,2-Dichloroethane Pyrolysis in a Laboratory Tubular Reactor, Ind. Eng. Chem. Res. 40 (2001) 2428–2436. doi:10.1021/ie0006460.

[9] O. Altin, S. Eser, Analysis of Solid Deposits from Thermal Stressing of a JP-8 Fuel on Different Tube Surfaces in a Flow Reactor, Ind. Eng. Chem. Res. 40 (2001) 596–603. doi:10.1021/ie0004491.

[10] S. Eser, R. Venkataraman, O. Altin, Deposition of Carbonaceous Solids on Different Substrates from Thermal Stressing of JP-8 and Jet A Fuels, Ind. Eng. Chem. Res. 45 (2006) 8946–8955. doi:10.1021/ie060968p.

[11] G. Liu, Y. Han, L. Wang, X. Zhang, Z. Mi, Solid Deposits from Thermal Stressing of n-Dodecane and Chinese RP-3 Jet Fuel in the Presence of Several Initiators, Energy Fuels. 23 (2009) 356–365. doi:10.1021/ef800657z.
[12] T. Edwards, Cracking and Deposition Behavior of Supercritical Hydrocarbon Aviation Fuels, Combustion Science and Technology. 178 (2006) 307–334. doi:10.1080/00102200500294346.
[13] B.M. Fabuss, J.O. Smith, C.N. Satterfield, Advances in Petroleum Chemistry Refining, Interscience, n.d.
[14] O. Altin, S. Eser, Carbon deposit formation from thermal stressing of petroleum fuels, Am. Chem. Soc. Div. Fuel Chem. 49 (2004).
[15] R.T.K. Baker, D.J.C. Yates, J.A. Dumesic, Filamentous Carbon Formation over Iron Surfaces, in: Coke Formation on Metal Surfaces, American Chemical Society, 1983: pp. 1–21. doi:10.1021/bk-1983-0202.ch001.
[16] F.D. Kopinke, G. Zimmermann, S. Nowak, On the mechanism of coke formation in steam cracking—conclusions from results obtained by tracer experiments, Carbon. 26 (1988) 117–124. doi:10.1016/0008-6223(88)90027-9.
[17] S. Shao, H. Zhang, Y. Wang, R. Xiao, L. Heng, D. Shen, Catalytic Pyrolysis of Biomass-Derived Compounds: Coking Kinetics and Formation Network, Energy Fuels. 29 (2015) 1751–1757. doi:10.1021/ef5026505.
[18] H.M. Jeong, M.W. Seo, S.M. Jeong, B.K. Na, S.J. Yoon, J.G. Lee, W.J. Lee, Pyrolysis kinetics of coking coal mixed with biomass under non-isothermal and isothermal conditions, Bioresource Technology. 155 (2014) 442–445. doi:10.1016/j.biortech.2014.01.005.
[19] F.D. Kopinke, G. Zimmermann, G.C. Reyniers, G.F. Froment, Relative rates of coke formation from hydrocarbons in steam cracking of naphtha. 2. Paraffins, naphthenes, mono-, di-, and cycloolefins, and acetylenes, Ind. Eng. Chem. Res. 32 (1993) 56–61. doi:10.1021/ie00013a009.
[20] J. Wang, M.-F. Reyniers, G.B. Marin, Influence of Dimethyl Disulfide on Coke Formation during Steam Cracking of Hydrocarbons, Ind. Eng. Chem. Res. 46 (2007) 4134–4148. doi:10.1021/ie061096u.
[21] N. Gascoin, G. Abraham, P. Gillard, Synthetic and jet fuels pyrolysis for cooling and combustion applications, Journal of Analytical and Applied Pyrolysis. 89 (2010) 294–306. doi:10.1016/j.jaap.2010.09.008.
[22] H. Huang, L.J. Spadaccini, D.R. Sobel, Fuel-Cooled Thermal Management for Advanced Aeroengines, J. Eng. Gas Turbines Power. 126 (2004) 284–293. doi:10.1115/1.1689361.
[23] H. Lander, A.C. Nixon, Endothermic fuels for hypersonic vehicles, Journal of Aircraft. 8 (1971) 200–207. doi:10.2514/3.44255.
[24] H. Huang, L. Spadaccini, D. Sobel, Endothermic Heat-Sink of Jet Fuels for Scramjet Cooling, in: 2002. doi:10.2514/6.2002-3871.
[25] R. Jiang, G. Liu, X. Zhang, Thermal Cracking of Hydrocarbon Aviation Fuels in Regenerative Cooling Microchannels, Energy Fuels. 27 (2013) 2563–2577. doi:10.1021/ef400367n.
[26] R. Bates, J. Edwards, M. Meyer, Heat Transfer and Deposition Behavior of Hydrocarbon Rocket Fuels, in: 41st Aerospace Sciences Meeting and Exhibit, American Institute of Aeronautics and Astronautics, n.d. doi:10.2514/6.2003-123.

[27] M.J. DeWitt, T. Edwards, L. Shafer, D. Brooks, R. Striebich, S.P. Bagley, M.J. Wornat, Effect of Aviation Fuel Type on Pyrolytic Reactivity and Deposition Propensity under Supercritical Conditions, Ind. Eng. Chem. Res. 50 (2011) 10434–10451. doi:10.1021/ie200257b.
[28] B. Stiegemeier, M. Meyer, R. Taghavi, A Thermal Stability and Heat Transfer Investigation of Five Hydrocarbon Fuels, in: 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, American Institute of Aeronautics and Astronautics, n.d. doi:10.2514/6.2002-3873.
[29] R. Venkataraman, S. Eser, Characterization of Solid Deposits Formed from Short Durations of Jet Fuel Degradation: Carbonaceous Solids, Ind. Eng. Chem. Res. 47 (2008) 9337–9350. doi:10.1021/ie8010066.
[30] A.R. Mohan, S. Eser, Analysis of Carbonaceous Solid Deposits from Thermal Oxidative Stressing of Jet-A Fuel on Iron- and Nickel-Based Alloy Surfaces, Ind. Eng. Chem. Res. 49 (2010) 2722–2730. doi:10.1021/ie901283r.
[31] S. Tang, N. Shi, J. Wang, A. Tang, Comparison of the anti-coking performance of CVD TiN, TiO2 and TiC coatings for hydrocarbon fuel pyrolysis, Ceramics International. 43 (2017) 3818–3823. doi:10.1016/j.ceramint.2016.12.036.
[32] B. Jin, K. Jing, J. Liu, X. Zhang, G. Liu, Pyrolysis and coking of endothermic hydrocarbon fuel in regenerative cooling channel under different pressures, Journal of Analytical and Applied Pyrolysis. 125 (2017) 117–126. doi:10.1016/j.jaap.2017.04.010.
[33] I.C. Lee, H.C. Ubanyionwu, Determination of sulfur contaminants in military jet fuels, Fuel. 87 (2008) 312–318. doi:10.1016/j.fuel.2007.05.010.
[34] M.L. Huber, E.W. Lemmon, T.J. Bruno, Effect of RP-1 Compositional Variability on Thermophysical Properties, Energy Fuels. 23 (2009) 5550–5555. doi:10.1021/ef900597q.
[35] L.S. Ott, A.B. Hadler, T.J. Bruno, Variability of The Rocket Propellants RP-1, RP-2, and TS-5: Application of a Composition- and Enthalpy-Explicit Distillation Curve Method, Ind. Eng. Chem. Res. 47 (2008) 9225–9233. doi:10.1021/ie800988u.
[36] MIL-DTL-25576 E PROPELLANT ROCKET GRADE KEROSENE, (n.d.). http://everyspec.com/MIL-SPECS/MIL-SPECS-MIL-DTL/MIL-DTL-25576E_16072/ (accessed October 31, 2017).
[37] B.C. Windom, T.J. Bruno, Assessment of the Composition and Distillation Properties of Thermally Stressed RP-1 and RP-2: Application to Fuel Regenerative Cooling, Energy Fuels. 25 (2011) 5200–5214. doi:10.1021/ef201077a.
[38] T.J. Fortin, T.J. Bruno, Assessment of the Thermophysical Properties of Thermally Stressed RP-1 and RP-2, Energy Fuels. 27 (2013) 2506–2514. doi:10.1021/ef400193d.
[39] R. Fernández-Varela, J.M. Andrade, S. Muniategui, D. Prada, F. Ramírez-Villalobos, The comparison of two heavy fuel oils in composition and weathering pattern, based on IR, GC-FID and GC–MS analyses: Application to the Prestige wreackage, Water Research. 43 (2009) 1015–1026. doi:10.1016/j.watres.2008.11.047.
[40] B.K. Lavine, J. Ritter, A.J. Moores, M. Wilson, A. Faruque, H.T. Mayfield, Source Identification of Underground Fuel Spills by Solid-Phase Microextraction/High-Resolution Gas Chromatography/Genetic Algorithms, Anal. Chem. 72 (2000) 423–431. doi:10.1021/ac9904967.
[41] J.S. Chickos, A.E. Wentz, D. Hillesheim-Cox, M.J. Zehe, Measurement of the Vaporization Enthalpy of Complex Mixtures by Correlation-Gas Chromatography. The

Vaporization Enthalpy of RJ-4, a High-Energy-Density Rocket Fuel at T = 298.15 K, Ind. Eng. Chem. Res. 42 (2003) 2874–2877. doi:10.1021/ie020920l.
[42] G. Liu, L. Wang, H. Qu, H. Shen, X. Zhang, S. Zhang, Z. Mi, Artificial neural network approaches on composition–property relationships of jet fuels based on GC–MS, Fuel. 86 (2007) 2551–2559. doi:10.1016/j.fuel.2007.02.023.
[43] M. Billingsley, T. Edwards, L. Shafer, T. Bruno, Extent and Impacts of Hydrocarbon Fuel Compositional Variability for Aerospace Propulsion Systems, in: 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, American Institute of Aeronautics and Astronautics, n.d. doi:10.2514/6.2010-6824.
[44] X. Sun, C.M. Zimmermann, G.P. Jackson, C.E. Bunker, P.B. Harrington, Classification of jet fuels by fuzzy rule-building expert systems applied to three-way data by fast gas chromatography—fast scanning quadrupole ion trap mass spectrometry, Talanta. 83 (2011) 1260–1268. doi:10.1016/j.talanta.2010.05.063.
[45] K.M. Pierce, S.P. Schale, Predicting percent composition of blends of biodiesel and conventional diesel using gas chromatography–mass spectrometry, comprehensive two-dimensional gas chromatography–mass spectrometry, and partial least squares analysis, Talanta. 83 (2011) 1254–1259. doi:10.1016/j.talanta.2010.07.084.
[46] B. Kehimkar, J.C. Hoggard, L.C. Marney, M.C. Billingsley, C.G. Fraga, T.J. Bruno, R.E. Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, Journal of Chromatography A. 1327 (2014) 132–140. doi:10.1016/j.chroma.2013.12.060.
[47] H. Lu, Q. Shi, J. Lu, G. Sheng, P. Peng, C.S. Hsu, Petroleum Sulfur Biomarkers Analyzed by Comprehensive Two-Dimensional Gas Chromatography Sulfur-Specific Detection and Mass Spectrometry, Energy Fuels. 27 (2013) 7245–7251. doi:10.1021/ef401239u.
[48] X. Gao, L. Pang, S. Zhu, W. Zhang, W. Dai, D. Li, S. He, Gas purge microsyringe extraction coupled to comprehensive two-dimensional gas chromatography for the characterization of petroleum migration, Organic Geochemistry. 106 (2017) 30–47. doi:10.1016/j.orggeochem.2017.01.007.
[49] A. Giri, M. Coutriade, A. Racaud, K. Okuda, J. Dane, R.B. Cody, J.-F. Focant, Molecular Characterization of Volatiles and Petrochemical Base Oils by Photo-Ionization GC×GC-TOF-MS, Anal. Chem. 89 (2017) 5395–5403. doi:10.1021/acs.analchem.7b00124.
[50] H. Potgieter, R. Bekker, J. Beigley, E. Rohwer, Analysis of oxidised heavy paraffininc products by high temperature comprehensive two-dimensional gas chromatography, Journal of Chromatography A. 1509 (2017) 123–131. doi:10.1016/j.chroma.2017.06.046.
[51] R.L. Webster, P.M. Rawson, C. Kulsing, D.J. Evans, P.J. Marriott, Investigation of the Thermal Oxidation of Conventional and Alternate Aviation Fuels with Comprehensive Two-Dimensional Gas Chromatography Accurate Mass Quadrupole Time-of-Flight Mass Spectrometry, Energy Fuels. 31 (2017) 4886–4894. doi:10.1021/acs.energyfuels.7b00178.
[52] B. Kehimkar, B.A. Parsons, J.C. Hoggard, M.C. Billingsley, T.J. Bruno, R.E. Synovec, Modeling RP-1 fuel advanced distillation data using comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry and partial least squares analysis, Anal Bioanal Chem. 407 (2015) 321–330. doi:10.1007/s00216-014-8233-6.

[53] B. Omais, M. Courtiade, N. Charon, D. Thiébaut, A. Quignard, M.-C. Hennion, Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate, Journal of Chromatography A. 1218 (2011) 3233–3240. doi:10.1016/j.chroma.2010.12.049.
[54] C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, Partial least squares analysis of rocket propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection, Talanta. 153 (2016) 203–210. doi:10.1016/j.talanta.2016.03.016.
[55] K.J. Johnson, B.J. Prazen, D.C. Young, R.E. Synovec, Quantification of naphthalenes in jet fuel with GC×GC/Tri-PLS and windowed rank minimization retention time alignment, J. Sep. Science. 27 (2004) 410–416. doi:10.1002/jssc.200301640.
[56] V. Abrahamsson, N. Ristic, K. Franz, K. Van Geem, Comprehensive two-dimensional gas chromatography in combination with pixel-based analysis for fouling tendency prediction, Journal of Chromatography A. 1501 (2017) 89–98. doi:10.1016/j.chroma.2017.04.021.
[57] J.R. Radović, K.V. Thomas, H. Parastar, S. Díez, R. Tauler, J.M. Bayona, Chemometrics-Assisted Effect-Directed Analysis of Crude and Refined Oil Using Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry, Environ. Sci. Technol. 48 (2014) 3074–3083. doi:10.1021/es404859m.
[58] P. de Peinder, T. Visser, R. Wagemans, J. Blomberg, H. Chaabani, F. Soulimani, B.M. Weckhuysen, Sulfur Speciation of Crude Oils by Partial Least Squares Regression Modeling of Their Infrared Spectra, Energy Fuels. 24 (2010) 557–562. doi:10.1021/ef900908p.
[59] C.M. Andersen, R. Bro, Variable selection in regression—a tutorial, J. Chemometrics. 24 (2010) 728–737. doi:10.1002/cem.1360.
[60] S. Wold, J. Trygg, A. Berglund, H. Antti, Some recent developments in PLS modeling, Chemometrics and Intelligent Laboratory Systems. 58 (2001) 131–150. doi:10.1016/S0169-7439(01)00156-3.
[61] W. Cai, Y. Li, X. Shao, A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra, Chemometrics and Intelligent Laboratory Systems. 90 (2008) 188–194. doi:10.1016/j.chemolab.2007.10.001.
[62] Å. Rinnan, M. Andersson, C. Ridder, S.B. Engelsen, Recursive weighted partial least squares (rPLS): an efficient variable selection method using PLS, J. Chemometrics. 28 (2014) 439–447. doi:10.1002/cem.2582.
[63] H. Kimura, J. Goto, S. Yasuda, S. Sakurai, M. Yumura, D.N. Futaba, K. Hata, Unexpectedly High Yield Carbon Nanotube Synthesis from Low-Activity Carbon Feedstocks at High Concentrations, ACS Nano. 7 (2013) 3150–3157. doi:10.1021/nn305513e.

Chapter 9. Enhancing the Chemical Selectivity in Discovery-based Analysis with Tandem Ionization Time-of-Flight Mass Spectrometry Detection for Comprehensive Two-Dimensional Gas Chromatography

This chapter was reproduced from C.E. Freye[†], N.R. Moore[†], R.E. Synovec, "Enhancing the Chemical Selectivity in Discovery-based Analysis with Tandem Ionization Time-of-Flight Mass Spectrometry Detection for Comprehensive Two-Dimensional Gas Chromatography" *Journal of Chromatography A* 1537 (2018), 99-108.

[†] These authors contributed equally to this work

9.1 INTRODUCTION

Comprehensive two-dimensional (2D) gas chromatography coupled with time-of-flight mass spectrometry (GC × GC – TOFMS) is a powerful instrumental platform for the analysis of complex mixtures such as those found in food chemistry [1–4], metabolomics [5–7], and petroleum products [8–10]. While a TOFMS is an attractive detector due to its fast data acquisition rate and full range mass sensitivity [3,11], identification of species is not always possible, or as confident as desired, due to similar fragmentation patterns between species that share similar structures. TOFMS and quadrupole MS generally employ electron impact (EI) ionization at 70 eV [8–12] when coupled with GC × GC separations. Limitations of hard ionization have brought about a demand for soft (lower) ionization methods (12–15 eV). Soft ionization can provide the preservation of the molecular ion and different fragmentation patterns, which is especially important for the identification of isomers [13,14]. There are numerous methods for soft ionization such as chemical ionization (CI) [15,16], field ionization (FI) [17,18], photoionization (PI) [19–21], and atmospheric pressure chemical ionization (APCI) [22–24]. However, all these methods

for soft ionization require a different ion-source which is time-consuming and resource intensive if a single TOFMS instrument is being used. Recently, TOFMS instrumentation for GC × GC has been introduced that provides a variable ionization source allowing for ionization energy to be tuned between 10 eV and 70 eV [13,14,25]. This variable ionization source has been used to study volatile organic compound profiles in human blood, increasing confidence in identification by providing more structurally significant mass fragments [14]. However, the multiple ionization energies were collected over several chromatographic runs which required the data to be aligned before any of the data processing could be performed. Recently this TOFMS instrument has been commercially modified to allow concurrent collection of two, user defined ionization energies, i.e., tandem ionization. Tandem ionization (TI) provides a hard ionization energy (70 eV) with a soft ionization energy (e.g. 14 eV), thus providing complementary streams of MS data without the need for retention time alignment between the two sets of data for each sample run. Fig. D.1 (Supporting Information) shows a representation of how tandem ionization data is collected by the instrument, in which mass spectra are collected at a rate of 50 Hz for each ionization energy.

Due to the complexity and enormous amount of information within a single GC × GC – TOFMS chromatogram, there is a need for data analysis strategies to uncover meaningful chemical information from complex samples. Meaningful chemical information is often buried in the background of less meaningful chemical signal and noise. Tandem ionization in conjunction with replicate injections will result in large data sets. The large data sets render identification of chemical features a time consuming, difficult process, especially when these chemical features are not known ahead of time, such as in non-targeted data mining studies that employ supervised comparison of samples/chromatograms [26,27]. A recent development to address many of these challenges in non-targeted discovery-based analysis of GC × GC – TOFMS data is tile-based

Fisher ratio (F-ratio) analysis [9,27–30], which is an ANOVA technique. F-ratio analysis prioritizes statistical significance in the class comparison of chemical analyte "features" over absolute signal. Indeed, the tile-based F-ratio software has been recently applied to yeast metabolites [29] and chemically altered diesel fuel [30] elucidating statistically significant class-distinguishing analytes in complex sample matrices. Recently, the software was optimized using receiver operator curves (ROC) to determine the ideal signal-to-noise (*S/N*) threshold and number of mass channels (*m*/*z*) used to calculate the average F-ratio in the generation of a hit list of chemical features [9]. A S/N of 10 and the top 10 *m*/*z* per hit were found to increase the true positive probability versus the false positive probability (i.e. increase discovery of true positives while decreasing false positives). While F-ratio analysis is amenable to discovering differences in GC × GC – TOFMS, identification and quantitation must be performed using a deconvolution technique. For this purpose, parallel factor analysis (PARAFAC) has been shown to be extremely effective for analyte deconvolution, identification, and quantification, making use of the third order advantage [31].

Herein, we investigate the complementary mass spectral information provided from the TI feature of the TOFMS using optimized non-targeted F-ratio analysis of GC × GC –TI – TOFMS data. Diesel fuel was spiked with twelve analytes. Diesel fuel was selected for this study as it provided a complex matrix background, while the spiked analytes were chosen to provide challenging cases for discovery, deconvolution, and identification. The complementary benefit of using a hard ionization energy (70 eV data set) concurrent with the soft ionization energy (14 eV data set) is explored, based upon the F-ratio hit lists produced. In addition, the two data sets are fused and F-ratio analysis is performed to further investigate the complementary benefits of the TI feature of the TOFMS. Combinatorial null distribution analysis for each hit list is applied,

providing a statistical threshold for the analyst [28–30]. After a "potential" hit location has been identified, PARAFAC is used to deconvolute the signal yielding quantitative information and identification, demonstrating the work flow from discovery to quantitative analysis.

9.2 EXPERIMENTAL

Diesel fuel (~1 gallon) was obtained from a local fueling station (Seattle, WA, USA). A non-native internal standard of 1,2,3-trichlorobenzene (0.5170 g), was spiked into one l of diesel fuel to create a stock solution of 608 ppm trichlorobenzene. This stock solution was used for all subsequent serial dilutions. A neat stock solution of twelve analyte compounds was made by adding ~45 mg each to a scintillation vial. The contents of the vial were quantitatively transferred to an amber bottle of known mass. The stock solution was then added until the final mass reached ~56 g, creating ~800 ppm concentration of the twelve spiked analytes. This solution was then serially diluted using the stock solution to concentration levels of ~50 ppm. A list of each spiked analyte, exact mass added, resulting spike concentration, purity and supplier is provided in Table D.1 (Supporting Information).

GC × GC – TI – TOFMS data was collected using an Agilent 7890A GC (Agilent Technologies, Palo Alto, CA, USA) coupled to a BenchTOF-Select (Markes International, Llantrisant, UK) equipped with a SepSolve Analytical flow modulator [32–35]. A reverse column configuration was employed with 30.0 m × 250 μm × 0.25 μm Rxi-17 Sil MS primary column (^{1}D) and a 3.5 m × 180 μm × 0.18 μm Rxi-1 MS as a secondary column (^{2}D) [8,9]. After the ^{2}D separation, the flows were split with a 1 m × 180 μm transfer line into TOFMS and a 0.74 m × 320 μm transfer line into the FID. This resulted in ~ 25% of the ^{2}D column effluent being split to the TOFMS and ~75% being split to the FID. The GC inlet and transfer line were set to 285 °C. Ultrahigh purity helium (Grade 5, 99.999%, Praxair, Seattle, WA, USA) was used as the

carrier gas at a constant flow of 1.8 ml/min on the ^{1}D column and 20 ml/min on the ^{2}D column. A 3.7 m × 100 µm of deactivated silica was used as a bleed line to balance the flows. An Agilent 7683 autosampler was used to inject 1 µl of diesel with a sample split at 200:1, resulting in ~62 pg of each spiked analyte reaching the TOFMS detector. Prior to injection, HPLC grade hexane and acetone, obtained from Fisher Scientific, were used as solvent rinses. The ^{1}D column was held at 40 °C for 3 min and then ramped at 5 °C/min to 285 °C and held for 2 min. The ^{2}D column was kept at the same nominal temperature as the ^{1}D column. The modulation period was 2.0 s with a 80 ms flush time. The effluent from the ^{1}D column was collected in a 25 µl sample loop before it was transferred to the ^{2}D column. The ion source was set to 285 °C. Mass channels (*m*/*z*) 35–350 were collected at 0.01 amu mass resolution at 50 spectra/s for both ionization energies (14 and 70 eV). The filament was held at 1.0 V then ramped to 1.6 V after 10 s and held constant for the rest of the acquisition. The FID temperature was set to 285 °C and data was collected at 100 Hz. The FID, with its higher data collection rate than the TOFMS, was used to assess the separation performance of the instrument, to ensure good ^{2}D chromatographic peak shapes were obtained. Six injection replicates each were obtained for the 50 ppm spike solution and the stock solution.

The GC × GC – TI – TOFMS data were baseline corrected using the instrument software (ChromSpace, version 1.0, Markes International, Llantrisant, UK). The baseline corrected data were converted to a netCDF file and was imported into Matlab 2015b (Mathworks, Inc., Natick, MA, USA) using an in-house algorithm. The mass spectral dimension was binned from 0.01 amu resolution to unit mass resolution after importation to reduce computation time with the current F-ratio software, while focusing on the benefits of the TI feature of the TOFMS. However, collecting the data at the higher mass resolution leaves open the possibility to explore software refinements in the future. The 50 ppm spike solution was compared to the stock solution using an in-house tile-

based F-ratio software, an algorithm that identifies differences between the two classes of samples, such as different concentrations of analytes. For a thorough a description of the software, readers are directed to previous publications [9,27–30]. For the F-ratio software executions, in the initial tiling step, the tile size on the ^{1}D dimension was 20 s and the ^{2}D dimension was 80 ms. All chromatograms were normalized to the internal standard 1,2,3-trichlorobenzene (*m*/*z* 180). The *S/N* threshold was set to 10 to initially calculate the F-ratio per tile (per *m*/*z*). At the 95% confidence level for the six spiked versus six unspiked diesel sample class comparison [36], a $F_{critical}$ (F_{crit}) threshold of 4.96 was then implemented to discard statistically insignificant F-ratios on a per tile (per *m*/*z*) basis, prior to calculating the average F-ratios. Implementation of the F_{crit} threshold is a new step relative to previous reports [9,27–30], as this step improved the ability to distinguish between true positives and false positives in the subsequent average F-ratio hit list. Next, on a per tile basis the average F-ratio was calculated. A maximum of 10 *m*/*z* (but at least 3 *m*/*z*) that passed the S/N threshold and F_{crit} threshold were required to calculate an average F-ratio per tile [9]. The pinning and clustering step was then automatically applied to remove redundant hit tiles. A 2D cluster window of 12 s on the ^{1}D dimension by 60 ms on the ^{2}D dimension was used. Finally, hit lists were calculated for both the 70 eV and 14 eV ionization energies. In addition, the data sets were fused (concatenated) by stacking the mass channels from two different ionization energies and the hit list was calculated. F-ratio calculations were performed identical to individual ionization energies. The workflow from the collection of the data, to the F-ratio analysis, to the eventual identification of the spiked analytes is provided in Fig. D.2 (Supporting Information).

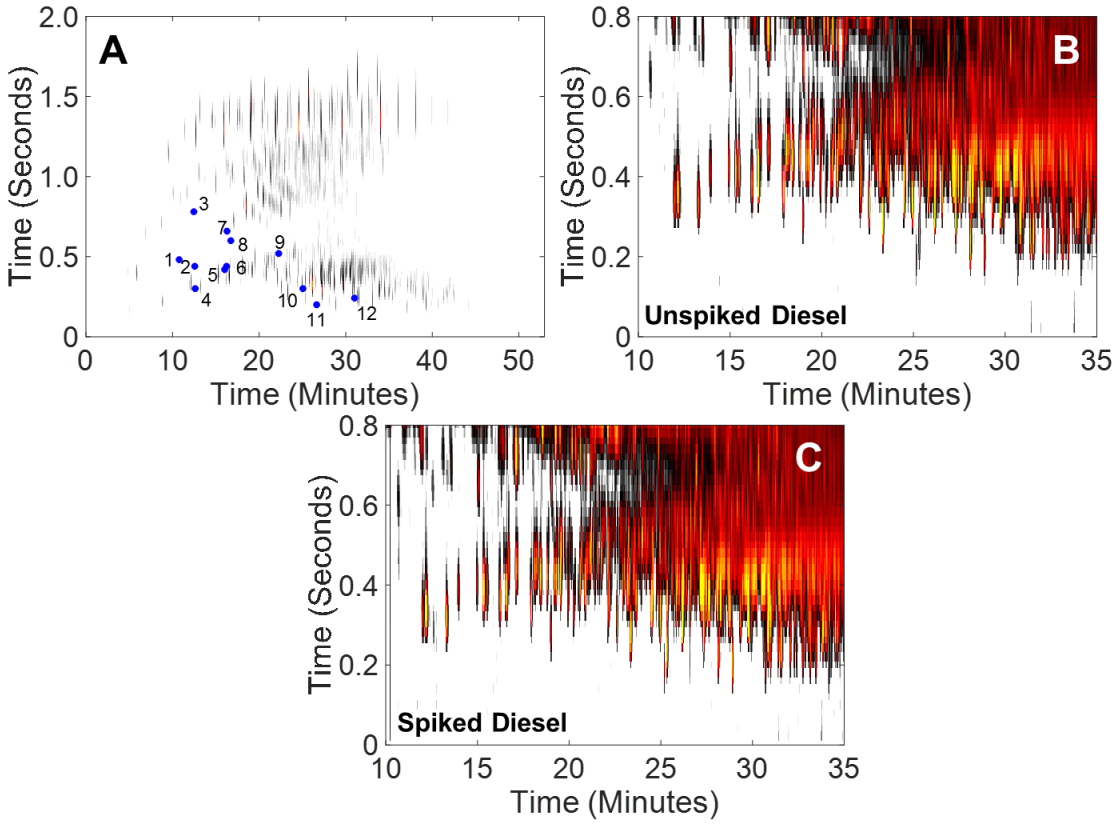


Figure 9.1 (A) Total ion current (TIC) GC × GC chromatogram of the spiked diesel at 70 eV annotated with the location of the twelve spiked analytes at a nominal concentration of 50 ppm. Table D.1 (Supporting Information) provides the identification and actual concentrations of each spiked analyte. (B) An enhanced portion of the unspiked diesel chromatogram at 70 eV has been provided and has been logarithmically scaled to, in principle, intensify analytes present at lower concentrations. (C) An enhanced portion of the spiked diesel chromatogram at 70 eV is shown and has been logarithmically scaled to intensify analytes present at lower concentrations.

9.3 RESULTS AND DISCUSSION

9.3.1 *GC × GC separations and fisher ratio analysis with tandem ionization TOFMS*

Using the GC × GC – TI– TOFMS in the reverse column configuration, an excellent 2D chromatographic separation was obtained. Fig. 9.1A shows the total ion current (TIC) chromatogram of the 70 eV data with the location of the twelve spiked analytes at the low nominal concentration of ~ 50 ppm. Table D.1 (Supporting Information) gives the names and actual concentrations of the twelve analytes. Ten of the analytes are non-native to this sample of diesel fuel while two are present at low concentration. Although the 14 and 70 eV chromatograms nominally look the same (see Fig. D.3, Supporting Information), there are key differences to be exploited in the F-ratio analysis. It would be a challenging endeavor for an analyst to find the differences between the spiked and unspiked TIC chromatograms at either ionization energy. Only through a priori knowledge could the analyst locate all twelve spiked analytes using specific m/z and inspecting the known 2D elution times. To illustrate this point, comparison of a spiked versus unspiked TIC chromatogram for the 70 eV ionization energy is provided in Fig. 9.1B and C, which have been logarithmically scaled to, in principle, enhance the signal due to the low concentration spiked analytes in the portion of the separation in which the spiked analytes predominate. Of the twelve spiked analytes, one is fully resolved on both dimensions (2,5-dimethylthiophene), three are partially resolved on the 2D dimension (1-chlorohexane, 1-ethylnapthalene, and limonene), and the remaining eight are moderately to severely overlapped. For example, a comparison of spiked versus unspiked TIC chromatographic data (70 eV) for 1-chlorohexane and 2-decanone is provided in Fig. 9.2. 1-Chlorohexane is readily observed in the TIC, while the 2-decanone cannot be observed due to the severe overlap from the diesel matrix background peaks.

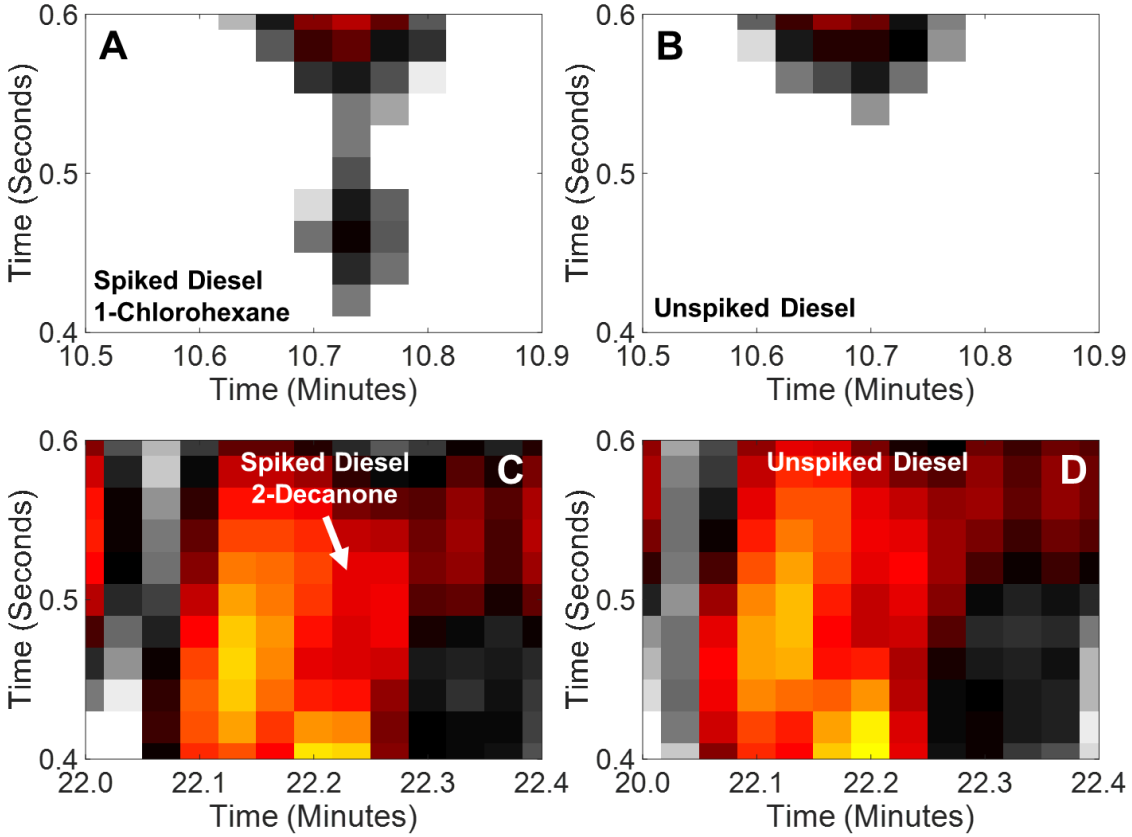


Figure 9.2 Portions of the spiked and unspiked diesel GC × GC chromatograms at 70 eV are shown for 1-chlorohexane and 2-decanone. When comparing the spiked and unspiked diesel chromatograms some minor retention time shifting on both ^{1}D and ^{2}D can be seen, which is readily handled by the tile-based F-ratio software (A) Portion of the spiked diesel GC × GC chromatogram containing 1-chlorohexane. (B) Portion of unspiked diesel GC × GC chromatogram, without 1-chlorohexane. (C) Portion of the spiked GC × GC chromatogram containing 2-decanone. (D) Portion of the unspiked GC × GC chromatogram, without 2-decanone.

F-ratio analysis provides a tremendous level of data reduction, as the algorithm marches through the various steps, culminating in the calculation of the average F-ratios in the hit list: tiling, S/N threshold, F_{crit} threshold, number of *m*/*z* required, and pinning and clustering for redundant hit removal. For both data sets there were 158,900 pixels (total spectra collected) in the 2D chromatograms (1589 ^{1}D modulations × 100 ^{2}D mass spectra) with 315 *m*/*z* per spectrum resulting in over 50 million possible F-ratios. Tiling each 2D chromatogram reduces the total number of possible hits to 1,244,250. For the 70 eV F-ratio comparison, application of the S/N and F_{crit} thresholds, followed by enforcing the number of *m*/*z* required coupled with calculation of the average F-ratio reduces the possible hits to 7470. Finally, redundant hit removal reduced the final hit list to 1713 average F-ratios for the 70 eV data set. Likewise, for the 14 eV F-ratio comparison the final hit list has 1771 entries. Finally, when the two data sets were fused, the final hit list has 1915 entries.

The average F-ratio distributions for the 70 eV, 14 eV, and fused data sets are provided in Fig. 9.3. None of the distributions have an average F-ratio <5 due to application of the F_{crit} threshold. The 70 eV and 14 eV distributions both have an apex at a F-ratio of ~10. Additionally, both data sets are sufficiently similar in shape and magnitude, thus fusing the data sets into a signal data set was deemed appropriate for the purpose of exploring the added chemical selectivity provided by tandem ionization coupled with F-ratio analysis. If the distributions were too dissimilar the analyst could still fuse them together, but the identified *m*/*z* would be biased towards a specific ionization energy. Null distribution analysis (details not shown for brevity) was performed for all three analyses [28–30], with results indicated in Fig. 9.3 insets. With the 70 eV data set, a F-ratio null threshold of 36 at a 0.1% false discovery rate (FDR) with a 95% confidence level was determined. Likewise, the F-ratio null threshold was 36 and 35 for the 14 eV data set

and the fused data set, respectively. Application of the F-ratio null threshold reduced the hit lists from 1713 entries to 9 entries for the 70 eV data set, 1771 to 9 for the 14 eV data set, and 1915 to 10 for the fused data set (70 and 14 eV). This corresponds to an overall data reduction to only ~0.0006% of the original GC × GC – TI – TOFMS data (see Fig. D.2). The analyst is recommended to evaluate hits slightly below the F-ratio null threshold in order to not overlook readily discovered true positives, as is practiced herein.

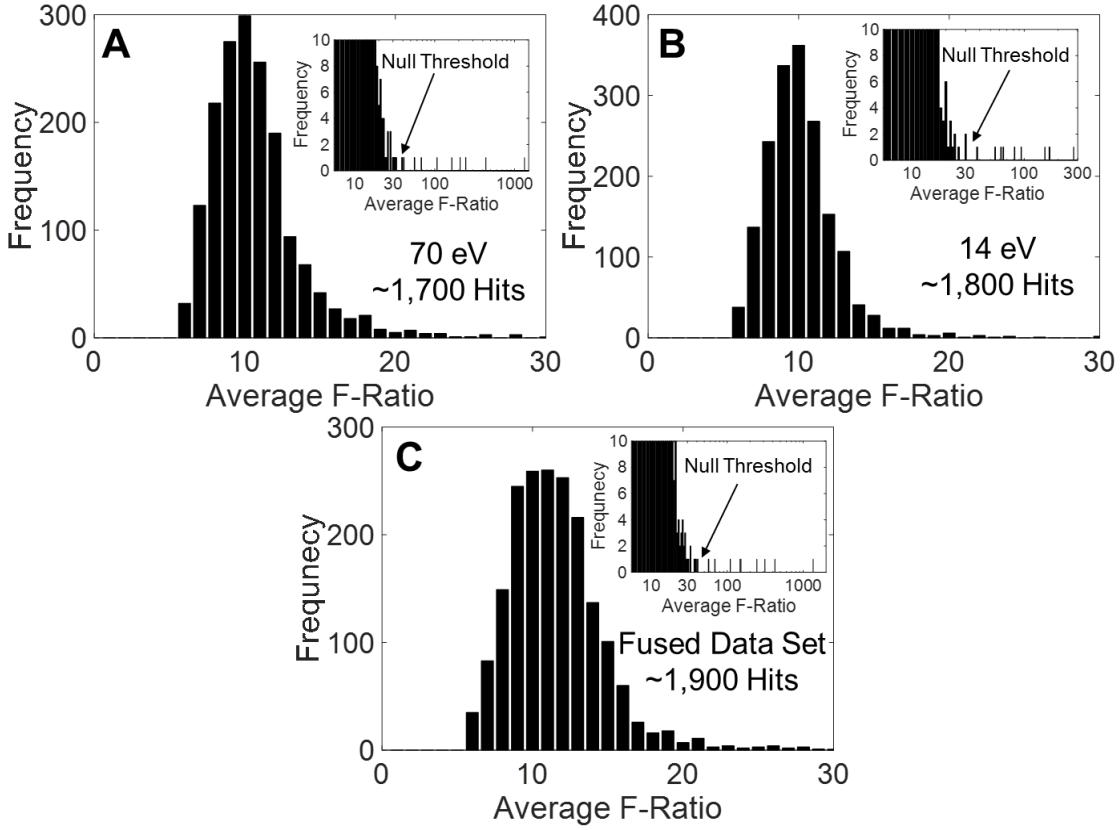


Figure 9.3 The average F-ratio distributions for the three comparisons of spiked versus unspiked diesel. Null distribution analysis was performed [28–30], with the F-ratio null threshold indicated in the inset figures, which provides an estimate of the 0.1% false discovery rate (FDR) at a 95% confidence level. (A) The 70 eV data set comparison, with a F-ratio null threshold of 36. (B) The 14 eV data set comparison, with a F-ratio null threshold of 36. (C) The comparison using the fused data set, combining the 70 eV and 14 eV data sets, with a F-ratio null threshold of 35.

A representative example of the complementary nature of the mass spectra for one of the twelve spiked analytes (butyrophenone), and how this results in complementary information in the F-ratio analysis is presented in Fig. 9.4. A "head-to-toe" plot comparing the 70 eV to the 14 eV mass spectra are provided in Fig. 9.4A. The 70 eV ionization energy results in more fragmentation while the 14 eV contains higher *m*/*z* fragments. The match value (MV) between the 70 and 14 eV spectra was calculated to provide a quantitative metric to infer complementariness of the two ionization energies for a given analyte. Calculated in this fashion, a lower MV indicates that the 70 and 14 eV spectra differ, while a higher MV (approaching 1000) indicates that the 70 and 14 eV spectra are similar. For butyrophenone this MV was 741, indicating there are significant spectral differences that the F-ratio analysis can likely leverage to provide more chemical selectivity. The "F-ratio spectra" for the 70 and 14 eV ionization energies are provided in Fig. 9.4B, where an F-ratio spectrum is the plot of the F-ratios as a function of *m*/*z*. The 70 eV F-ratio analysis identified three lower m/z that were not present in the 14 eV. In addition, the *m*/*z* identified when the two ionization energies data sets were fused are starred (★). The "head-to-toe" mass spectra plots with the 70–14 eV mass spectra MV, and the F-ratio spectra comparing the 70 eV to 14 eV for the other eleven analytes is provided in Fig. D.4 (Supporting Information).

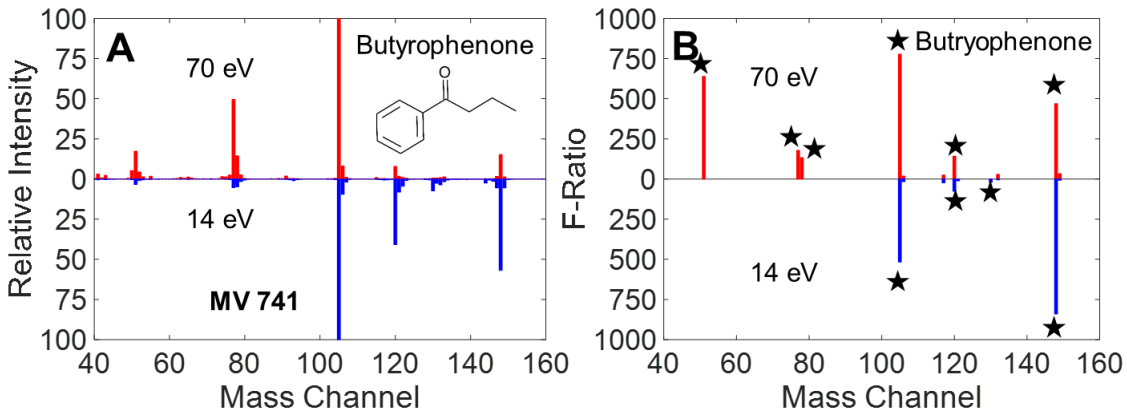


Figure 9.4 (A) Head-to-toe plots comparing the 70 eV to 14 eV mass spectra for butyrophenone. The match value (MV) between the 70 and 14 eV spectra is indicated, providing a quantitative metric to infer complementariness of the mass spectra at the two ionization energies. (B) Head-to-toe plots comparing the 70–14 eV F-ratio spectra for butyrophenone. The mass channels, m/z, starred () are those found when the two data sets (70 eV and 14 eV) are fused prior to performing the F-ratio analysis. Head-to-toe plots comparing the 70–14 eV mass spectra and the F-ratio spectra for the other eleven analytes can be found in Fig. D.4 (Supporting Information).

9.3.2 *Fisher ratio analysis hit list: 70 eV hard ionization*

A summary of the 70 eV F-ratio hit list is provided in Table 9.1 with ^{1}D retention time ($^{1}t_{R}$), ^{2}D retention time ($^{2}t_{R}$), average F-ratio value, and the qualifying m/z in order of largest to smallest (left to right) used to calculate the average F-ratio. Eleven of the twelve analytes were discovered with two false positives found interspersed toward the bottom of this abbreviated hit list. One of the analytes, 2-heptanol was not discovered due to the S/N and Fcrit threshold. 2-Heptanol only has one major fragment (m/z 45) which is the only m/z that passed the S/N threshold that is associated with 2-heptanol. The F-ratio for m/z 45 at the tile that contained 2-heptanol had a value of 0.74 which does not pass the F_{crit} threshold. The analyte 1-ethylnapthalene is natively

present at a concentration of 38.2 ppm [9], so with the spike concentration of 40.9 ppm the resulting concentration ratio of 1.9 was more challenging than the other spiked (non-native) analytes, resulting in a lower F-ratio. Based upon the null distribution analysis, see Fig. 9.3A, at the determined F-ratio null threshold of 36 (0.1% FDR), nine of the twelve spiked analytes are above the F-ratio null threshold with one false positive. Note that hits #12 (2-decanone) and #13 (1-ethylnaphthalene) are just below the 0.1% FDR and would be readily discovered in a practical application of the F-ratio analysis. In the 70 eV hit list in Table 9.1, a total of 94 *m/z* passed the various criteria to be included in the average F-ratio calculation. The *m/z* only found within in the 70 eV hit list (in comparison to the 14 eV hit list in Table 9.2) are **bolded** (34 m/z or 36.2% of the 94 m/z), and false positive m/z are in parentheses (8 m/z or 8.5%). Generally, the unique (**bolded**) m/z are smaller m/z, because hard ionization results in more fragmentation compared to the soft ionization which emphasizes higher m/z as shown in Table 9.2. Proportionately more false

Name	$^{1}t_R$ (min)	$^{2}t_R$ (s)	Hit #	Average F-ratio	Mass Channels, *m/z* (top 10 F-ratios, minimum of 3)
2,5-dimethylthiophene	12.63	0.28	1	1332	112,111,97,113,114,**59**,**52**,**71**,99,**50**
1-chlorohexane	10.73	0.48	2	437	93,56,91,**41**,55,43
butyrophenone	26.63	0.20	3	245	105,**51**,148,**77**,120,**78**,149,132,117,106
5-decyne	16.73	0.60	4	207	54,67,81,68,**55**,82,**53**,95,109,96
t-butylbenzene	16.23	0.44	5	163	119,134,93,**99**,135,**72**,**91**,**92**
limonene	16.30	0.66	6	107	107,136,92,93,121,94,**122**,78,108,**68**
3-octanone	16.03	0.42	7	67	72,99, (134)
cyclooctane	12.57	0.66	8	55	56,84,**41**,55,83,70,**42**,112,69,**43**
1-nonanol	25.03	0.30	9	41	**93**,**77**,**127**,**64**,**63**, (133,142,104,131,130)
2-decanone	22.23	0.52	12	32	43,58,156,59,**72**, (80,137)
1-ethylnapthalene	31.03	0.24	13	31	**58**,**70**,**101**,**113**,141,**89**,**114**,**87**,**126**,**71**
2-heptanol	12.47	0.78	—[a]	—[a]	—[a]

Table 9.1 Summary of the 70 eV hit list. Eleven of the 12 analytes were found with two false positives (hits #10 and #11). The *m/z* **bolded** are only found within the 70 eV hit list and those in parentheses are false positives. In the final hit list, the null distribution threshold for a 0.1% false discovery rate (FDR) at a 95% confidence interval is an average F-ratio of 36. (a) 2-Heptanol was not found due to being below the *S/N* threshold.

positive m/z are found further down the hit list as one would expect for this challenging study in which the spike level of ~ 50 ppm is approaching the limit of detection (LOD) for the separation conditions implemented.

9.3.3 *Fisher ratio analysis hit list: 14 eV soft ionization*

Table 9.2 summarizes the 14 eV ionization energy F-ratio hit list for the twelve spiked analytes. 1-nonanol (hit # 8 in the 70 eV comparison) and 2-heptanol were not discovered due to being below the S/N threshold. At the determined F-ratio null threshold of 36, see Fig. 9.3B, nine of the twelve spiked analytes are above the F-ratio null threshold with one false positive. In the 14 eV hit list, a total of 81 m/z passed the various criteria to be included in the average F-ratio calculation. The m/z only found in the 14 eV hit list are **bolded** (21 m/z or 25.9% of the 81 m/z),

Name	^1t_R (min)	^2t_R (s)	Hit #	Average F-ratio	Mass Channels, *m/z* (top 10 F-ratios, minimum of 3)
2,5-dimethylthiophene	12.63	0.28	1	276	112,114,113,97,**98**,111,**61**,99
butyrophenone	26.63	0.20	2	169	148,105,120,117,**130**,106,**121**,149,132
1-chlorohexane	10.73	0.48	3	166	91,55,93,**42**,56,43,**69**,**84**,**92**
5-decyne	16.73	0.60	4	153	81,54,109,67,95,68,82,96
t-butylbenzene	16.23	0.44	5	94	119,135,134,93,**118**, (126)
cyclooctane	12.57	0.66	6	81	55,84,112,83,56,70,**57**,69,**54**,**68**
limonene	16.30	0.66	7	64	**91**,121,107,92,136,93,**80**,94,108,78
2-decanone	22.23	0.52	8	61	59,58,156,43, (137,152,326,78)
3-octanone	16.03	0.42	9	55	72,**128**,99,**73**, (134,135)
1-ethylnaphthalene	31.03	0.24	293	12	141**,142,154,156,152,143,**(191)
1-nonanol	25.03	0.30	—[a]	—[a]	—[a]
2-heptanol	12.47	0.78	—[b]	—[b]	—[b]

Table 9.2 Summary of the 14 eV hit list. The top 9 hits correspond to 9 of 12 the analytes. 1-ethylnapthalene was found, albeit at a much lower hit number. The m/z **bolded** are only found within the 14 eV hit list and those in parentheses are false positives. In the final hit list, the null distribution threshold for a 0.1% FDR is an average F-ratio of 36. (a) 1-nonanol, and (b) 2-heptanol were not found due to being below the S/N threshold.

and false positive *m*/*z* are in parentheses (8 m/z or 9.9%). In contrast to the 70 eV results in Table 9.1, the **bolded** *m*/*z* are generally larger m/z due to the lower ionization energy.

9.3.4 *Fisher ratio analysis of fused data sets: 70 eV and 14 eV*

					Mass Channels, *m/z* (top 10 F-ratios, min of 3)	
Name	**^1t_R (min)**	**^2t_R (s)**	**Hit #**	**Average F-ratio**	**70 eV**	**14 eV**
2,5-dimethylthiophene	12.63	0.30	1	1353	112,111,**97**,**113**,**114**,**59**,**52**,**71**	111,112
1-chlorohexane	10.73	0.48	2	425	**42**,93,**56**,**91**,**41**,55,**43**	55,93
butyrophenone	26.63	0.20	3	313	105,**51**,148,**77**,120,**78**	148,120,105,**132**
5-decyne	16.73	0.60	4	246	54,67,81,**68**,**55**	**109**,54,81,**95**,67
t-butylbenzene	16.23	0.44	5	152	119,134,93,**99**,135,**72**	119,134,135,93
limonene	16.30	0.66	6	147	107,**136**,92,**93**,121,**94**	**91**,121,107,92
cyclooctane	12.57	0.66	7	110	**56**,84,**41**,55,**83**,**70**,**42**	55,84,**112**
2-decanone	22.23	0.52	8	68	**43**,58,156,59,**72,** (80)	59,58,156, (137)
3-octanone	16.03	0.42	9	56	72,99	72,**128**,99,**73**, (134,135)
1-nonanol	25.03	0.30	10	41	**93**,**77**,**127**,**64**,**63**,(133,142,104,131,130)	—[a]
1-ethylnaphthalene	31.03	0.24	14	33	**58**,**70**,**101**,**113**,141,**89**,**114**,**87**	141,**142**
2-heptanol	12.47	0.78	—[b]	—[b]	—[b]	—[b]

Table 9.3 Summary of the F-ratio hit list for the fused 70 eV and 14 eV data sets. The top 10 hits correspond to 10 of the analytes with 1-ethylnaphthalene being found at hit number 14. The m/z **bolded** are only found within one of the ionization voltages and those in parentheses are false positives. Compared to the individual F-ratio analyses (Tables 1 and 2), complementary m/z are emphasized and there are significantly fewer false positive m/z. The null distribution threshold for a 0.1% FDR is an average F-ratio of 35. (a) 1-nonanol has no m/z that contribute at 14 eV due being below the S/N threshold, and (b) 2-heptanol was not discovered due to being below the S/N threshold.

Finally, in contrast to Tables 9.1 and 9.2, in which the two different ionization energy data sets were analyzed independently, Table 9.3 provides the F-ratio analysis when the two ionization energy data sets have been fused into a single data set. Evidence indicating the selectivity advantage of fusing the data sets is provided. Eleven of the twelve analytes were discovered along

with three false positives above the F-ratio null threshold of 35, per Fig. 9.2C. 2-Heptanol was not found due to being below the S/N and F_{crit} thresholds and 1-nonanol did not have any *m*/*z* from the 14 eV due to being below the S/N threshold. It is noteworthy that when the data sets were fused, there are significantly fewer false positive *m*/*z*. In general, the true positive hits were more confidently "discovered" in the hit list, further away from the F-ratio null threshold. Overall, the TI feature of the TOFMS enhances the F-ratio analysis as indicated by the complementary information expressed in Tables 9.1–3. Indeed, by fusing the two data sets there were 107 *m*/*z* that passed the various criteria of the F-ratio software with a total of 49 unique *m*/*z* (45.8% unique *m*/*z*), and only 9 false positive *m*/*z* (8.4% of the 107 *m*/*z*). Granted the 70 eV data set contributed more unique ions when the two data sets were fused, but this is to be expected due to increased fragmentation.

To further investigate the complementary information provided by tandem ionization, the MV between the 70 eV and 14 eV mass spectra were calculated based on the in-house library spectra (see Fig. D.4, Supporting Information). The average MV for all twelve analytes was 777 ± 80. In principle, the lower the MV between the 70 and 14 eV spectra, the greater likelihood that there will be unique m/z identified in the F-ratio calculations of the fused data set. However, several analytes are overlapped significantly by large interferences so it is not possible to directly study this metric for all twelve analytes. Four analytes were sufficiently resolved to consider this metric: 1-ethylnapthalene, 2,5-dimethylthiophene, limonene, and 1-chlorohexane, with 70–14 eV MV of 638, 666, 720, and 882, respectively. Correspondingly, the number of unique m/z identified for 1-ethylnapthalene, 2,5-dimethylthiophene, limonene, and 1-chlorohexane are 8, 6, 4, and 5, respectively, indicating there may be a cause/effect relationship between the unique number m/z and the MV. More investigation into this topic is warranted.

9.3.5 *Analyte discovery and identification*

A familiar figure-of-merit dealing with analyte detection is the limit of detection (LOD) [37,38]. The most common definition for the LOD, defined by the IUPAC in 1975, is when the analyte signal is three times the standard deviation of the baseline noise [39]. Less familiar yet critical figures-of-merit for analyte detection relate to analyte discovery (e.g., in an F-ratio analysis), and analyte identification. Herein, we refer to these figures-of-merit as the "limit of discovery" and "limit of identification". These limits are determined in the context of a given experimental scenario and data set, and are highly dependent on the matrix background and refer to the notion that there is a minimum analyte concentration in which analyte discovery and identification can be confidently accomplished. The basic idea is, for a given analyte, the same concentration likely does not define these two limits. Indeed, the limit of discovery is often lower than the limit of identification (as demonstrated below), especially if identification requires a complete mass spectrum which can be more challenging to obtain when major interferents are present. Thus "discovering" a statistically significant change in analyte from one sample class to the next does not guarantee it will be identified.

For the F-ratio analysis, the limit of discovery is defined by the F-ratio null threshold. As previously mentioned, the hit lists were investigated to an F-ratio value right below the F-ratio null threshold (Fig. 9.3). Besides providing a F-ratio threshold, the null distribution analysis yields a quantitative metric for the FDR at a given confidence level (i.e. a 0.1% FDR at a 95% confidence). F-ratio analysis locates statistically significant signal differences between two sample classes (i.e. "discover" chemical differences), but F-ratio analysis does not directly provide analyte identification. The 2D locations indicated by F-ratio analysis must be "mined" to identify (and quantify) the analytes that are responsible for the statistical differences. PARAFAC is utilized to

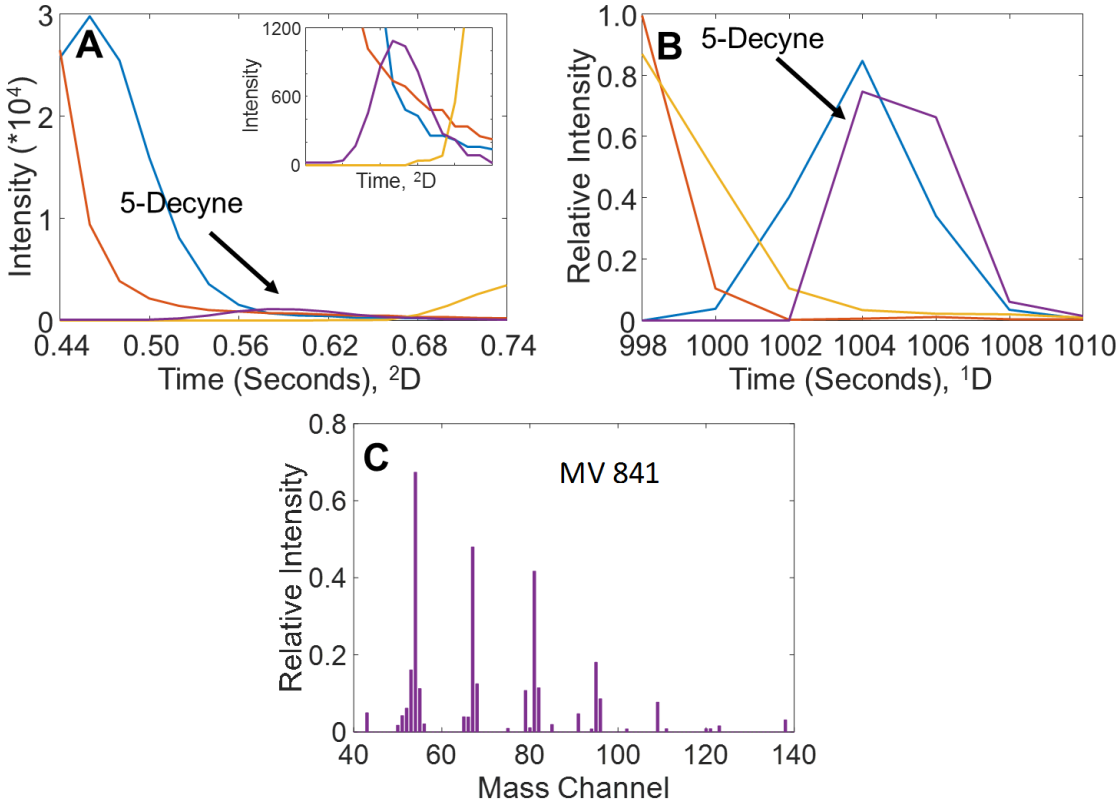


Figure 9.5 PARAFAC deconvolution of 5-decyne from the 70 eV data set, overlapped by three interfering analytes. (A) ^{2}D chromatographic profile which contains the relative analyte concentration information. (B) ^{1}D chromatographic profile (C) Mass spectrum and match value (MV). In this instance, in contrast to Fig. 9.4, the MV is between the 70 eV spectrum deconvoluted by PARAFAC relative to the 70 eV library spectrum. The PARAFAC deconvolution of 5-decyne at 14 eV can be found in Fig. D.5 (Supporting Information).

"mine" these locations with the aim to obtain a full mass spectrum for a given F-ratio hit. PARAFAC decomposes the 3-way data array (GC × GC – TOFMS) into a series of components in which each analyte is described by a single component. Each component is given by the outer product of ^{1}D and ^{2}D chromatographic and mass spectrometric vectors. PARAFAC has a trilinearity requirement which may be interrupted in GC × GC – TOFMS due to modulation-to-

modulation ^2t_R shifting [40]. The potential loss of trilinearity can be readily mitigated by using an appropriate P_M (1–2 s) in conjunction with an appropriate temperature ramp rate (5–10 °C/min). PARAFAC, as well as F-ratio analysis, does not require the chromatographic peak shapes to be Gaussian(s) on either/both the ^{1}D or ^{2}D separations. Herein the limit of identification is defined as a MV > 800, for the MV of the deconvoluted spectrum relative to the in-house spectrum for a given ionization energy.

The center of the 2D location of each hit is defined by the cluster step of the F-ratio software, so each analyte of interest is centered in a 2D window for PARAFAC with dimensions of 12 s (6 modulations) on ^{1}D by 300 ms on ^{2}D. This 2D window for PARAFAC is extracted from the chromatogram in order to reduce computation time and reduce the number of factors required in the model. When performing the deconvolution, an analyst could use the F-ratio spectrum to help guide their deconvolution through the use of soft constraints or visually comparing the F-ratio spectrum to the mass spectrum. Fig. 9.5 shows the PARAFAC deconvolution of 5-decyne from the 70 eV data set, which was hit #4 in the 70 eV F-ratio comparison. Even though 5-decyne is overlapped by two large interfering compounds, PARAFAC deconvolution provides a high mass spectral MV of 841 (the traditional sample spectrum to library spectrum MV at a given ionization energy in this instance). PARAFAC deconvolution of 5-decyne at 14 eV can be found in Fig. D.5 (Supporting Information) with similar results. Table 9.4 provides a summary of the PARAFAC deconvolution for all twelve analytes at 70 eV and 14 eV. For comparison, the mass spectra were taken at the peak apex and then matched. While some analytes such as 2,5-dimethylthiophene are resolved on both GC × GC dimensions, the PARAFAC MV is much higher than the peak apex MV because PARAFAC removes noise, which results in a higher MV. Several spiked analytes, 3-octanone, 2-decanone, and 1-nonanol were not successfully deconvoluted and thus not

"identified." All three of these spiked analytes were not deconvoluted due to a lack of enough selective *m*/*z* and/or severe overlap with large interferences. 2-Heptanol was not "discovered" in the F-ratio analysis and so a deconvolution could not be performed. The complementary benefit of having both sets of ionization energy data is further exemplified in Table 9.4, where we note that the MV obtained for the 14 eV data are consistently higher than those obtained with the 70 eV data set. It appears whereas the 70 eV data set contributes more to discovering the spiked analytes (see Table 9.3), identification of the spiked analytes is more confident with the 14 eV data set. Moreover, the extra ionization energy mass spectra provide increased confidence of identification

		70 eV			14 eV		
Name	Average F-ratio, fused data set	Apex MV	PARAFAC MV	MV %RSD	Apex MV	PARAFAC MV	MV %RSD
2,5-dimethylhtiophene	1353	843	883	3.6	828	921	5.9
1-chlorohexane	425	792	898	5.5	742	922	1.5
butyrophenone	313	716	831	2.5	677	829	1.9
5-decyne	246	518	817	5.3	513	898	2.0
t-butylbenzene	152	667	830	1.4	519	874	2.9
limonene	147	824	857	3.7	896	928	3.5
cyclooctane	110	778	851	7.9	812	916	2.9
2-decanone	68	44	—[a]	—[a]	17	—[a]	—[a]
3-octanone	56	213	—[b]	—[b]	12	629	7.8
1-nonanol	41	13	—[a]	—[a]	—[c]	—[c]	—[c]
1-ethylnapthalene	33	807	821	0.3	896	900	3.0
2-heptanol	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]	—[d]

Table 9.4 The average F-ratios for the fused data set (combining the 70 eV and 14 eV data sets), followed by PARAFAC deconvolution to provide the mass spectrum of the spiked analytes. (a) 2-decanone and 1-nonanol were not deconvoluted due to severe interference. (b) 3-octanone was not identified in the 70 eV data due to not enough selective m/z. (c) 1-nonanol was not found with the 14 eV F-ratio analysis. (d) 2-heptanol was not discovered with F-ratio analysis, due to low S/N.

of an analyte by providing another MV.

We turn our attention to 2-decanone in order to discuss this distinction between the limit of discovery versus limit of identification. Fig. 9.6A shows a single modulation which corresponds to the highest intensity sub-peak of 2-decanone for the 70 eV data set. Visually it appears that 2-decanone is overlapped by four analytes. Fig. 9.6B shows the mass spectrum from the peak apex corresponding to the location of 2-decanone. The most intense *m*/*z* of 2-decanone (*m*/*z* 58) is ~10-fold less intense than the most intense *m*/*z* at this retention time. Essentially, most of the spectral features in the mass spectrum in Fig. 9.6B are due to the interferences. The corresponding figures for the 14 eV data set can be found in Fig. D.6 (Supporting Information). Fortunately, F-ratio analysis is sensitive to small class-to-class changes in concentration [27], enabling the software to

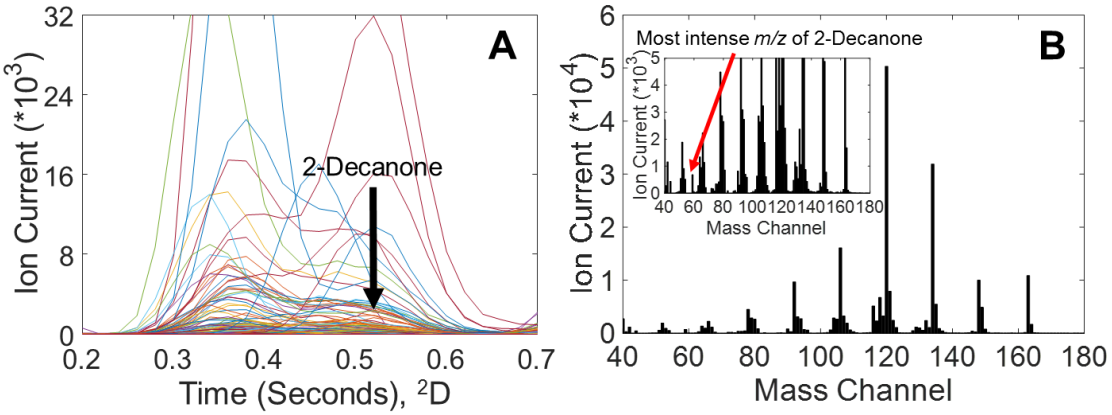


Figure 9.6 (A) From the 70 eV data set, a portion of the GC × GC data along the ^{2}D dimension for a single modulation is shown which corresponds to the highest intensity for 2-decanone which elutes at a ^2t_R of 0.52 s. Visually, it appears that at least four interfering analytes overlap. (B) Mass spectrum from the elution time of 2-decanone. The most intense mass channel (*m/z* 58) is ~10-fold less intense than the most intense m/z. While F-ratio analysis is able to discover the location of 2-decanone, PARAFAC deconvolution of the full mass spectrum was not successful. The corresponding figures for the 14 eV data set can be found in Fig. D.6 (Supporting Information).

discover extremely minor changes despite a large matrix background. Specially, 2-decanone was hit #12 in the 70 eV data set with 5 selective *m*/*z* and only 2 false positive *m*/*z* (Table 9.1), and even better at hit #8 in the 14 eV data set with 4 selective *m*/*z* and 4 false positive *m*/*z* (Table 9.2). Furthermore, fusing the two data sets improves the discovery of 2-decanone as it is hit #8 with 2 *m*/*z* unique to the 70 eV data set and 3 *m*/*z* in common between the two data sets, so a total of 8 selective *m*/*z*, and only 2 false positive *m*/*z* (Table 9.3). However, for both ionization energies, 2-decanone was not successfully deconvoluted for the current purpose of providing its full mass spectrum, hence it is below the limit of identification since an adequate MV could not be provided (Table 9.4).

9.4 CONCLUSION

Herein we have investigated the complementary information provided by tandem ionization time-of-flight mass spectrometry (TI – TOFMS) when coupled with GC × GC for discovery-based analysis. For this purpose, we sought to "discover" twelve analytes that were spiked into diesel fuel at low concentration. Investigating the hit list of the 70 eV and 14 eV ionization energies, and the fusion of these two data sets, resulted in the discovery of eleven of the twelve analytes. Due to the complementary nature of tandem ionization, the F-ratio spectra *m*/*z* identified by all of the hit lists aided in the discovery of analytes. Using PARAFAC deconvolution, nine of the twelve analytes were accurately identified based upon a 2D retention time match and mass spectral MV. Finally, discussion about the limit of discovery and limit of identification indicates that tandem ionization is a beneficial in aiding the analysis of complex samples such as diesel. It has been demonstrated how an analyst could implement this workflow for discovery based analysis.

9.5 SUPPORTING INFORMATION

Figures D.1 to D.6 and Table D.1 can be found in the Supporting Information which is available free of charge on the internet at

https://doi.org/10.1016/j.chroma.2018.01.008

The Supporting Information is also reprinted in Appendix D.

9.6 ACKNOWLEDGEMENTS

We thank SepSOLVE for loaning the instrumentation for this study. We specifically thank Pete Grosshans, Nick Bukowski, Wade Bontempo, Matt Edwards, Ian Biggs, and Ollie Stacey for helpful discussions and assistance with facilitating setting up the instrumentation.

9.7 REFERENCES

[1] D. Ryan, R. Shellie, P. Tranchida, A. Casilli, L. Mondello, P. Marriott, Analysis of roasted coffee bean volatiles by using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, J. Chromatogr. A 1054 (2004) 57-65.

[2] F. Magagna, A. Guglielmetti, E. Liberto, S.E. Reichenbach, E. Allegrucci, G. Gobino, C. Bicchi, C. Cordero, Comprehensive chemical fingerprinting of high-quality cocoa at early stages of processing: effectiveness of combined untargeted and targeted approaches for classification and discrimination, J. Agric. Food Chem. 65 (2017) 6329-6341.

[3] R. Shellie, P.J. Marriott, P. Morrison, Concepts and preliminary observations on the triple-dimensional analysis of complex volatile samples by using GC×GC−TOFMS, Anal. Chem. 73 (2001) 1336-1344.

[4] P.-H. Stefanuto, K.A. Perrault, L.M. Dubois, B. L'Homme, C. Allen, C. Loughnane, N. Ochiai, J.-F. Focant, Advanced method optimization for volatile aroma profiling of beer using two-dimensional gas chromatography time-of-flight mass spectrometry, J. Chromatogr. A 1507 (2017) 45-52.

[5] M.K. Das, S.C. Bishwal, A. Das, D. Dabral, A. Varshney, V.K. Badireddy, R. Nanda, Investigation of gender-specific exhaled breath volatome in humans by GC × GC – TOF-MS, Anal. Chem. 86 (2014) 1229-1237.

[6] W. Welthagen, R.A. Shellie, J. Spranger, M. Ristow, R. Zimmermann, O. Fiehn, Comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC × GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice, Metabolomics 1 (2005) 65-73.

[7] X. Huang, F.E. Regnier, Differential metabolomics using stable isotope labeling and two-dimensional gas chromatography with time-of-flight mass spectrometry, Anal. Chem. 80 (2008) 107-114.

[8] B. Omais, M. Courtiade, N. Charon, D. Thiébaut, A. Quignard, M.C. Hennion, Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate, J. Chromatogr. A 1218 (2011) 3233-3240.
[9] B.C. Reaser, B.W. Wright, R.E. Synovec, Using receiver operating characteristic curves to optimize discovery-based software with comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, Anal. Chem. 89 (2017) 3606-3612.
[10] M.K. Jennerwein, M. Eschner, T. Gröger, T. Wilharm, R. Zimmermann, Complete group-type quantification of petroleum middle distillates based on comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC×GC-TOFMS) and Visual Basic Scripting, Energy Fuels 28 (2014) 5670-5681.
[11] J. Dallüge, L.L.P. Van Stee, X. Xu, J. Williams, J. Beens, R.J.J. Vreuls, U.A.Th Brinkman, Unravelling the composition of very complex samples by comprehensive gas chromatography coupled to time-of-flight mass spectrometry: Cigarette smoke, J. Chromatogr. A 974 (2002) 169-184.
[12] J.F. Hamilton, P.J. Webb, A.C. Lewis, J.R. Hopkins, S. Smith, P. Davy, Partially oxidised organic components in urban aerosol using GCXGC-TOF/MS, Atmos. Chem. Phys. 4 (2004) 1279-1290.
[13] M.S. Alam, C. Stark, R.M. Harrison, Using variable ionization energy time-of-flight mass spectrometry with comprehensive GC×GC to identify isomeric species, Anal. Chem. 88 (2016) 4211-4220.
[14] L.M. Dubois, K.A. Perrault, P.H. Stefanuto, S. Koschinski, M. Edwards, L. McGregor, J.F. Focant, Thermal desorption comprehensive two-dimensional gas chromatography coupled to variable-energy electron ionization time-of-flight mass spectrometry for monitoring subtle changes in volatile organic compound profiles of human blood, J. Chromatogr. A 1501 (2017) 117-127.
[15] B. Munson, Chemical ionization mass spectrometry: ten years later, Anal. Chem. 49 (1977) 772A-775A.
[16] M.S.B. Munson, F.H. Field, Chemical ionization mass spectrometry. I. General Introduction, J. Am. Chem. Soc. 88 (1966) 2621-2630.
[17] J. Chen, N. Mclean, D. Hager, Prediction of molecular weight by-boiling-point distribution of middle distillates from gas chromatography−field ionization mass spectrometry (GC−FIMS), Energy Fuels 25 (2011) 719-726.
[18] W. Genuit, H. Chaabani, Comprehensive two-dimensional gas chromatography-field ionization time-of-flight mass spectrometry (GCxGC-FI-TOFMS) for detailed hydrocarbon middle distillate analysis, International Journal of Mass Spectrometry 413 (2017) 27-32.
[19] A. Giri, M. Coutriade, A. Racaudi, K. Okuda, J. Dane, R.B. Cody, J.F. Focant, Molecular characterization of volatiles and petrochemical base oils by photo-ionization GC × GC-TOF – MS, Anal. Chem. 89 (2017) 5394-5403.
[20] S. Mitschke, W. Welthagen, R. Zimmermann, Comprehensive gas chromatography – time-of-flight mass spectrometry using soft and selective photoionization techniques, Anal. Chem. 78 (2006) 6364-6375.
[21] L. Hanley, R. Zimmermann, Light and molecular ions: the emergence of vacuum UV single-photon ionization in MS, Anal. Chem. 81 (2009) 4174−4182.

[22] V.V. Lobodin, E.V. Maksimova, R.P. Rodgers, Gas chromatography/atmospheric pressure chemical ionization tandem mass spectrometry for fingerprinting the Macondo oil spill, Anal. Chem. 88 (2016) 6914-6922.
[23] T. Portolés, C. Sales, B. Gómara, J.V. Sancho, J. Beltrán, L. Herrero, M.J. González, F. Hernández, Novel analytical approach for brominated flame retardants based on the use of gas chromatography-atmospheric pressure chemical ionization-tandem mass spectrometry with emphasis in highly brominated congeners, Anal. Chem. 87 (2015) 9892-9899.
[24] A. Ballesteros-Gómez, J. de Boer, P.E.G. Leonards, Novel analytical methods for flame retardants and plasticizers based on gas chromatography, comprehensive two dimensional gas chromatography, and direct probe coupled to atmospheric pressure chemical ionization-high resolution time-of-flight-mass spectrometry, Anal. Chem. 85 (2013) 9572-9580.
[25] Markes International, Ltd. Application Note 537 (2016).
[26] K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, J. Chromatogr. A 1255 (2012) 3-11.
[27] L.C. Marney, W.C. Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher-ratio software for improved feature selection analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Talanta 115 (2013) 887-895.
[28] B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher ratio analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC–TOFMS) data using a null distribution approach, Anal. Chem. 87 (2015) 3812-3819.
[29] N.E. Watson, B.A. Parsons, R.E. Synovec, Performance evaluation of tile-based Fisher ratio analysis using a benchmark yeast metabolome dataset, J. Chromatogr. A 1459 (2016) 101-111.
[30] B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, Chemical characterization of the acid alteration of diesel fuel: Non-targeted analysis by two-dimensional gas chromatography coupled with time-of-flight mass spectrometry with tile-based Fisher ratio and combinatorial threshold determination, J. Chromatogr. A 1440 (2016) 179-190.
[31] J.C. Hoggard, R.E. Synovec, Parallel factor analysis (PARAFAC) of target analytes in GC × GC-TOFMS data: Automated selection of a model with an appropriate number of factors, Anal. Chem. 79 (2007) 1611-1619.
[32] M. Poliak, M. Kochman, A. Amiray, Pulsed flow modulation comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1186 (2008) 189–195.
[33] Q. Gu, F. David, F. Lynen, K. Rumpel, G. Xu, P. Vos, P. Sandra, Analysis of bacterial fatty acids by flow modulated comprehensive two-dimensional gas chromatography with parallel flame ionization detector/mass spectrometry, J. Chromatogr. A 1217 (2010) 4448-4453.
[34] P.A. Bueno, J.V. Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1027 (2004) 3-10.
[35] J.F. Griffith, W.L. Winniford, K. Sun, J. Luong, A reversed-flow differential flow modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1226 (2011) 116-123.

[36] R.E. Walpole, R.H. Myers, Probability and statistics for engineers and scientists, second ed., MacMillan Publishing Co., Inc., New York, 1978.
[37] K.H. Chang, Limit of detection and its establishment in analytical chemistry, Health Env. J. 2 (2011) 38-43.
[38] J. Krupčík, P. Májek, R. Gorovenko, J. Blaško, R. Kubinec, P. Sandra, Considerations on the determination of the limit of detection and the limit of quantification in one-dimensional and comprehensive two-dimensional gas chromatography, J. Chromatogr. A 1396 (2015) 117-130.
[39] G.L. Long, I.D. Winefordner, Limit of detection a closer look at the IUPAC definition, Anal. Chem. 55 (1983) 712A-724A.
[40] D.K. Pinkerton, B.A. Parsons, T.J. Anderson, R.E. Synovec, Trilinearity deviation ratio: a new metric for chemometric analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Anal. Chim. Acta 871 (2015) 66-76.

Chapter 10. Expanding the Limits of Detection and Quantitation in Chromatography through Integration

10.1 INTRODUCTION

The ability to detect a trace element or molecule is termed the limit of detection (LOD) which describes the lowest concentration or mass that can be measured. This value must be statistically different from the analytical blank but various approaches of measuring this value can lead to different values which could easily span an order of magnitude [1-6]. Indeed, many different organization define LOD in various terms. The US Enivronmental Protection Agency (EPA) uses method detection limit as the minimum concentration of an analyte that can be identified and measured at 99% confidence while the International Conference for Harmonization (ICH) defines LOD as the lowest amount of analyte in a sample which can be detected but not necessarily quantitated as an exact value [6]. The most commonly used definition of the LOD, which was defined by the IUPAC in 1975, is three times the standard deviation of the noise [7-12]. At this location, there is a 0.13% chance that the measured signal will fall within the distribution of the noise [12]. Thus, there is a chance, albeit small, that the measured signal is not a true measurement.

Limit of quantitation (LOQ) is the lowest concentration at which an analyte can be detected reliably for some predefined goal of bias and imprecision. At the LOQ there is an insignificantly small chance that the measured value will fall within the distribution of the noise [4-7]. Furthermore, fluctuations of the baseline will affect the analyte response signal by less than 20% [13]. The IUPAC definition of LOQ is ten times the standard deviation of the noise [7]. The original IUPAC definition was chosen for spectrochemical analysis (i.e. UV-Vis absorbance, ICP-AES), which does not account for chromatographic conditions where measurement of an analyte

is a time dependent response (e.g., spacial phenomenon observed in the time domain) [14]. Nevertheless, most chromatographic analysis for determining the LOQ (or LOD) use the IUPAC definition [6, 14].

Quantitation in chromatography is normally performed by determining the peak area of the analyte of interest. Perhaps the most common, current method used for determining the peak area in chromatography is the two-step method [15]. As the name implies, there are two "steps" that must be performed in order to obtain the peak area. The first "step," is to determine the *S/N* of the peak of interest. Briefly, the signal is determined by finding the signal at the peak maximum and dividing by the noise which is determined by finding the standard deviation for a portion of the chromatogram where there are no analytes present or column bleed. If the *S/N* of the peak is above the LOQ, the peak is integrated in the second "step." Another method proposed by Synovec and Yeung involved integrating the peak and then determining if the measured area was above the LOQ [16]. They proposed using the peak height to determine if the analyte was above the LOD (or LOQ) was inappropriate because much of the "peak" information was not included. Furthermore, measurement of analyte in chromatography is not univariate measurement. By integrating the peak of interest, the signal detectability is enhanced relative the original chromatogram.

Integration methods, such as cumulative sum, have been used to determine underlying trends which may not be apparent in raw data. Extracting signal from background noise is a problem in analysis; however, significant changes can be detected in the data from its background noise as shown by Page using the cumulative sum (CUSUM) test [17]. Essentially, CUSUM is a statistical method to detect the serial-dependence of the signal in the presence of background noise. The cumulative sum of random data points should fluctuate around a mean while detection of an

analyte will cause the trend line to drift away from the mean indicating a non-random process is occurring (i.e. response of a detector to an analyte). Using this technique, statistical detection methods such as CUSUM-charts have been developed to determine the changes from the mean level of a series of sequential data points [18-20]. CUSUM-charts have often been used in areas of quality control [21-23] and medical fields [24-26]. They have also been used in analytical nuclear chemistry to monitor background radioactivity counts in order to obtain more accurate measurements [27]. Additionally, this technique has been used to obtain information about the internal energy distribution of fluorescing nitrogen dioxide [28]. CUSUM-charts have been used in gas chromatography for quality control to ensure that the detector, columns, and other variables in the instrument are remaining constant in order to obtain reproducible results [29, 30].

Herein we report on an integration method to improve upon the *S/N* of peaks in a chromatographic system in an attempt to glean more accurate and precise quantitation of analytes. Converting the analyte response into a function of time through integration can result in an improved *S/N*. Integration of the chromatographic data converts the univariate data into a function of time which takes into account the peak shape which can be affected negatively by the separation media. This method requires a new definition to measure the signal as well as noise in order to determine the appropriate *S/N*. We have proposed a method to measure both these values and studied this method in a rigorous manner. Comparing integration of the analyte response and noise, we have identified the optimum manner to gain the maximum improvement for the *S/N*. Using the optimized conditions, we have applied this technique to both one and two dimensional separations. Additionally we have compared our technique to a commonly used method to quantitate chromatographic data, the two-step method, in order to evaluate the accuracy and precision of our technique. Finally, the integration method is herein demonstrated on experimental data obtained

from two dimensional gas chromatography with a time-of-flight mass spectrometer (GC × GC – TOFMS).

10.2 THEORY

A Gaussian function is often used to model chromatographic data and peaks can be described by,

$$S_i(t) = \frac{V_i R_i}{\sigma_i (2\pi)^{1/2}} exp(\frac{-(t-t_{R,i})^2}{2{\sigma_i}^2}) \qquad (10.1)$$

where $S_i(t)$ is the height of the peak at time t, defined by the standard deviation of the peak (σ_i), retention time (t_R) at the maximum $S_i(t)$, volume fraction injected (V_i), and analyte detection response factor (R_i). Integration of equation 10.1 will yield the area of the Gaussian function.

$$S_{i,AREA} = \int_{-\infty}^{\infty} \frac{V_i R_i}{\sigma_i (2\pi)^{1/2}} exp(\frac{-(t-t_{R,i})^2}{2{\sigma_i}^2})\ dt \qquad (10.2)$$

In practice, integration of data is normally done by performing a Riemann sum. A Riemann sum is calculated by dividing the region into shapes (most commonly squares for Gaussian distributions) and summing the area of those shapes. However, because the data is continuous and sampled at discrete time intervals, the $S_{i,\,AREA}$ is approximated by taking the Riemann sum of chromatographic peak (i.e. summing each data point). If data was collected at an infinitely fast rate, the Riemann sum for a peak would converge to the integral shown in equation 10.2. Because V_i and R_i are both independent of time, these variables can be factored out from equation 10.1. The remaining portion of equation 10.1, is referred to as a normal (or Gaussian) distribution which integrates to 1. Thus we can simplify the equation to:

$$S_{i,AREA} = V_i R_i \qquad (10.3)$$

By substituting $t = t_R$ into equation 10.1, the maximum of the Gaussian function can be calculated.

$$S_{i,MAX} = \frac{V_i R_i}{\sigma_i (2\pi)^{1/2}} \tag{10.4}$$

Traditionally, the maximum signal ($S_{i,MAX}$) is used to determine if the response of the analyte is above the LOD or the LOQ. If the response is below the LOD or LOQ, the peak is left out from the analysis (i.e. not identified/quantified). The improvement in the signal using the area compared to the maximum height of the function is given as:

$$S_{IMPROVEMENT} = \frac{S_{i,AREA}}{S_{i,MAX}} = \sigma_i (2\pi)^{1/2} \tag{10.5}$$

By using the peak area to determine if an analyte is above the LOD or LOQ, a significant improvement in the sensitivity can be gained. For a peak width (W_b) of 20 data points which is common in gas chromatography (GC), the signal is improved by a roughly a factor of 12.5 if the entire area of the peak is captured. However, equation 10.5 assumes that there is an infinite signal to noise (*S/N*).

Noise for chromatographic detectors often follows a normal distribution in which there is a uniform probability distribution over some interval [3]. The probability density of the normal distribution for the noise can be modeled by:

$$f(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{10.6}$$

Here μ is the mean of the distribution which can be set to zero through baseline correction of the chromatogram. The standard deviation is signified by σ. Measurement of the noise in a chromatographic system is typically done in a region where no analytes or column bleed is present. The standard deviation of a significant number of data points is used to define the noise of the chromatogram. When the noise is integrated the variance of the integrated domain can be summarized as:

$$\sigma_{Int}^2 = \sum_{i=1}^{n} \sigma_i^2 \tag{10.7}$$

where σ_{Int}^2 is the variance of the noise in the integrated domain, *n* is the number of points and σ_i is the standard deviation of the noise in the chromatographic system. Because we assume that the distribution of the noise is constant throughout the chromatogram, we can simply equation 10.7 to:

$$\sigma_{Int}^2 = n\sigma_0^2 \qquad (10.8)$$

Thus the standard deviation of the noise in the integrated domain is

$$\sigma_{Int} = \sqrt{n}\sigma_0 \qquad (10.9)$$

Therefore the increase in the noise from the traditional domain to the integrated is

$$Noise_{INCREASE} = \frac{\sigma_{Int}}{\sigma_0} = \sqrt{n} \qquad (10.10)$$

Assuming a W_b (4σ) of 20 data points and integrating across (6σ) to capture 99.7% of the peak area, the noise in the integrated domain will increase by a factor of $\sqrt{30}$ or roughly 5.5 in comparison to the tradition definition of noise.

Given we know the improvement in the signal from the traditional chromatogram to the integrated chromatogram as well as the increase in the noise, we can calculate a *S/N* improvement factor by combining equation 10.5 and 10.10.

$$(S/N)_{IMF} = \frac{Signal_{IMPROVEMENT}}{Noise_{INCREASE}} = \frac{\sigma_i(2\pi)^{1/2}}{\sqrt{n}} \qquad (10.11)$$

For a W_b of 20 data points the *S/N* should increase by a factor of 2.28 to the integrated domain from the traditional domain wherein σ_i is 5 and *n* is 30 (because of integrating across 6σ). As the number of data points to define a peak increases, the improvement factor will become higher increasing the *S/N* in the integrated domain compared to the traditional chromatographic domain.

10.3 EXPERIMENTAL

Chromatographic simulation, quantitation, and statistical calculations were all done in MATLAB version 2012b (MathWorks, Natick, MA). The software was written in-house. Gaussian noise was generated with a standard deviation of 0.003757 with a mean of zero. The standard deviation of the noise was determined by finding the average of the noise of ten chromatograms that were free of analytes and column bleed. These chromatograms were collected on an Agilent 6890 GC (Agilent Technologies, Palo Alto, CA, USA). For the chromatographic simulations, the W_b was 20 data points wide unless otherwise noted. Figure 10.1 shows the traditional means that *S/N* were measured as well as how the integrated signal and integrated noise were determined. Signal for the Gaussian peaks was determined by height above baseline of the peak center which was specified in the simulations. Thus *S/N* for the Gaussian peaks was calculated by taking the signal at peak center and dividing by the noise (which was specified). The entire peak was then integrated from -8σ to 8σ making sure the entire peak was captured. Signal was calculated for the integrated peak by finding the average of σ points (5 data points for a W_b of 20 data points) at a specified difference from the peak center. For example, to capture 6σ of the peak (for a 20 data point wide peak), the average of 5 points was found centered at -3σ and 3σ. The difference of the two values was used to find the signal of the integrated peak. Noise was calculated in the same manner.

Chromatograms of a test mixture were collected in order to demonstrate this technique on non-simulated data. The GC × GC – TOFMS system was an Agilent 6890N gas chromatograph equipped with an Agilent 7683 autoinjector coupled with a LECO Pegasus III TOFMS equipped with a 4D thermal modulator upgrade (LECO, St. Joseph, MI). The electron energy for the TOFMS was set to -70 eV with a detector voltage of 1600 V, and the data was collected at 100 Hz from 35

amu to 300 amu at unit mass resolution. The inlet was set to 275 °C. The oven was held at 40 °C for 1 minute, ramped at 5 °C/min to 300 °C and held for 2 minutes. There was a 5 °C offset on the secondary oven. The flow rate was 2.0 mL/min of ultra-high purity helium and the modulation period was 2.0 seconds. The first-dimension (^{1}D) column was a 30 meter × 250 µm inner diameter (I.D.) × 0.25 µm film thickness (D.F.) Rxi-17SilMS (diphenyl dimethyl polysiloxane) and the second dimension column (^{2}D) column was a 1.5 meter × 180 µm I.D. × 0.18 µm D.F. Rxi-1ms (dimethyl polysiloxane).

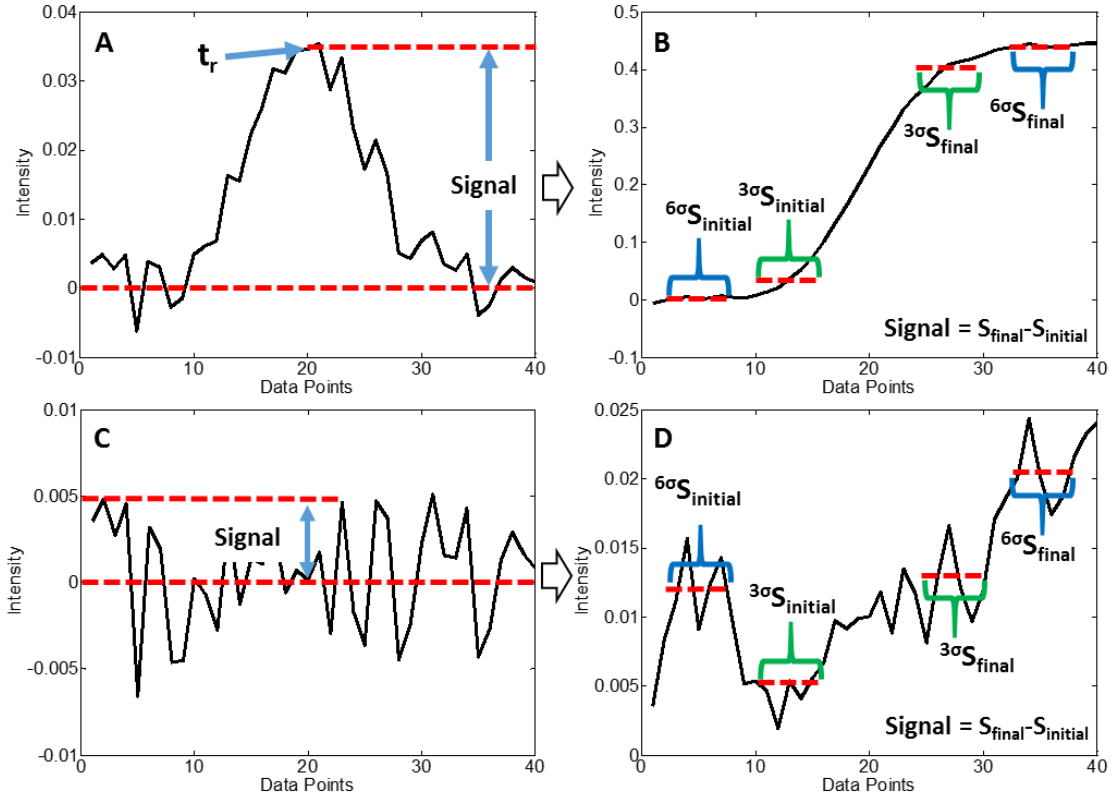


Figure 10.1 (A) Chromatographic peak as it elutes is shown. Signal is measured from baseline to maximum on the peak (often denoted as the retention time, t_R) (B) The chromatographic peak has been integrated and the signal has been transformed from univariate to integrated wherein the signal is dependent on time and incorporates the concentration profile (i.e. peak shape). Signal is defined as the difference of the averages of a section prior to the peak eluting and a section after the peak has eluted (C) Signal for noise has often been defined as height above baseline for a single pixel response (D) By integrating the noise, it is amplified and is subjected to a random walk effect. In order to determine the signal of the noise, the integrated noise is averaged prior to where the peak eluted and post peak elution. The difference of these values yields a signal for the noise in the integrated domain

10.4 RESULTS AND DISCUSSION

Traditionally, chromatographic peaks have been measured at their peak maximum on a pixel based level and this value has been used to determine if the peak is above the LOQ (or LOD). If the peak is above the LOQ, the concentration can be determined by finding the peak area. However, the concentration of the peak is a function of time due to its Gaussian shape and should not be measured in a pixel based manner because that only indicates the concentration of the peak at that specific point in time. Any asymmetry of the peak brought about by chromatographic interactions, will broaden the peak and lower the peak maximum. By integrating the peak, the signal can be transformed from univariate to integrated wherein the signal is dependent of time and incorporates the concentration profile (i.e. peak shape). The integrated shape is affected by the original chromatographic peak shape but the signal of the integrated peak is not affected by the peak shape unlike traditional means of measuring chromatographic peaks. Additionally, measuring the area of the peak through integration results in a higher signal. As shown in equation 10.2, the area of the peak is $\sigma(2\pi)^{1/2}$ times the peak height. For a W_b of 20 data points (4σ), the signal increases by a factor of roughly 12.5 if the entire peak area is captured.

Noise in detectors such as flame ionization (FID), time of flight mass spectrometry (TOFMS), or optical detectors (i.e. absorbance, florescence) exhibit Gaussian white noise which mathematically defined is a uniform probability distribution over a specific interval. The noise of a detector is typically measured by taking the standard deviation of a section of a chromatogram which is free of analytes and column bleed. Measurement of the noise in the integrated domain cannot be done similar to traditional way of determining noise because the noise has been transformed into a function with respect to time. The standard deviation of the noise in the integrated domain is different than that of the traditional means of measuring noise as a function

of the square root of the number of data points (equation 10.7). As seen in Figure 10.2A, the noise in the integrated domain is subjected to a random walk effect due to the random distribution of the noise. In order to determine the noise in the integrated domain, 1000 randomly generated noise files that were 40 data points long were integrated. This corresponds to -4σ to 4σ for a peak that is 20 data points wide. A subset of those curves can be seen in Figure 10.2A. The first five data points (1σ) were averaged and subtracted from the average of the last five data points which corresponds to 7σ of the peak area captured if a peak with W_b of 20 data points was integrated. Figure 10.2B shows the distribution that was generated from the 1000 replicates when 7σ is measured. The standard deviation of the distribution was 0.0215 which was identified as the noise of the integrated domain when 7σ of the peak is captured. The noise in the integrated domain has increased roughly by $\sqrt{33}$ which matches the theoretical prediction in equation 10.7.

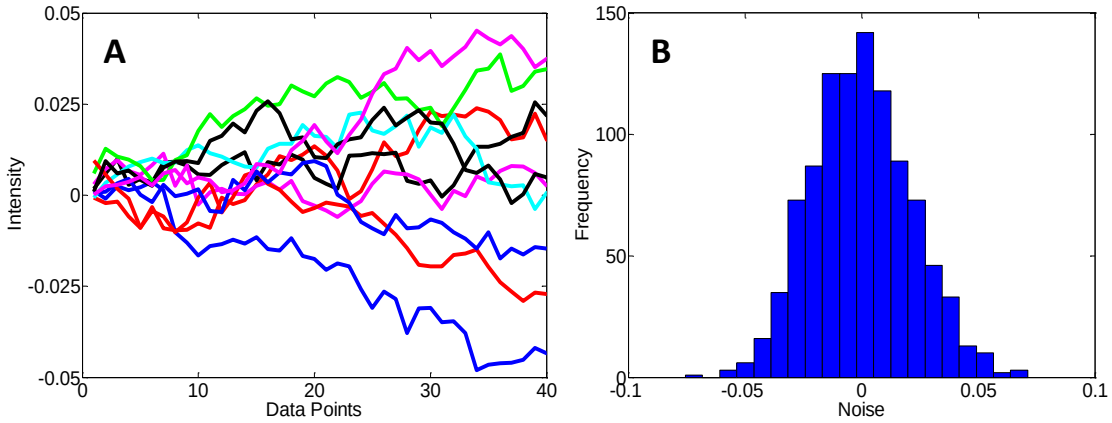


Figure 10.2 (A) Random noise was generated and integrated. As shown in Figure 10.1, signal is defined as the difference of an average of an area prior to peak elution and post peak elution. In A the first five data points were averaged and subtracted from the average of the last five data points. The baseline drifts as a function of the number of points integrated. (B) Distribution of the signal of the integrated random noise. This distribution was produced using 1000 replicates and has a standard deviation of 0.0215.

Integrating a lower number of points will result in a small distribution of the noise but conversely this would result in capturing less of the area underneath the peak. Figure 10.3A shows the increase in the noise dependent on the W_b and σ captured. Similar to Figure 10.2, noise was generated using a Gaussian distribution and integrated. For each W_b, varying amounts of σ (2σ to 7σ) were captured. The integrated noise was divided by the traditional definition of noise to find the increase in noise. As W_b increases, the noise increases by a square root of the number of points integrated as defined in equation 9 (i.e. the more points in a peak, the greater the noise). Capturing more area under the peak (i.e. 7σ compared to 2σ) results in a higher noise increase because the distribution of the noise. For a W_b of 20 data points, if you only integrate and capture 2σ of the peak, the noise will increase by a factor of roughly 2.5 from the original noise. However, if you integrate and capture 7σ of the peak, the noise will increase by a factor of roughly 6. Inverse to the noise increase, the signal increases as more area under the peak is captured and as W_b increases. Capturing 7σ results in 99.7% of the peak area being integrated while capturing 2σ results in only 68.2% of the peak area being captured as shown in Figure 10.3B. Thus the signal of the integrated peak will have a higher increase in signal if more of the peak is captured. Dividing the increase in the signal (S_{int}/S_0) by the noise (N_{int}/N_0) yields an improvement in the *S/N*. Figure 10.3C shows the improvement in signal to noise for various W_b and σ captured. For a W_b of 20 data points, the maximum improvement of signal is when 3σ of the peak is captured. However, 4σ captured is likely a better choice because if asymmetry is present in the peak, the larger integration area will capture the potential variation in the peak shape and is less than a 10% difference between the optimum of 3σ captured. This increase in *S/N* will lower the LOD and LOQ by a factor of 3.

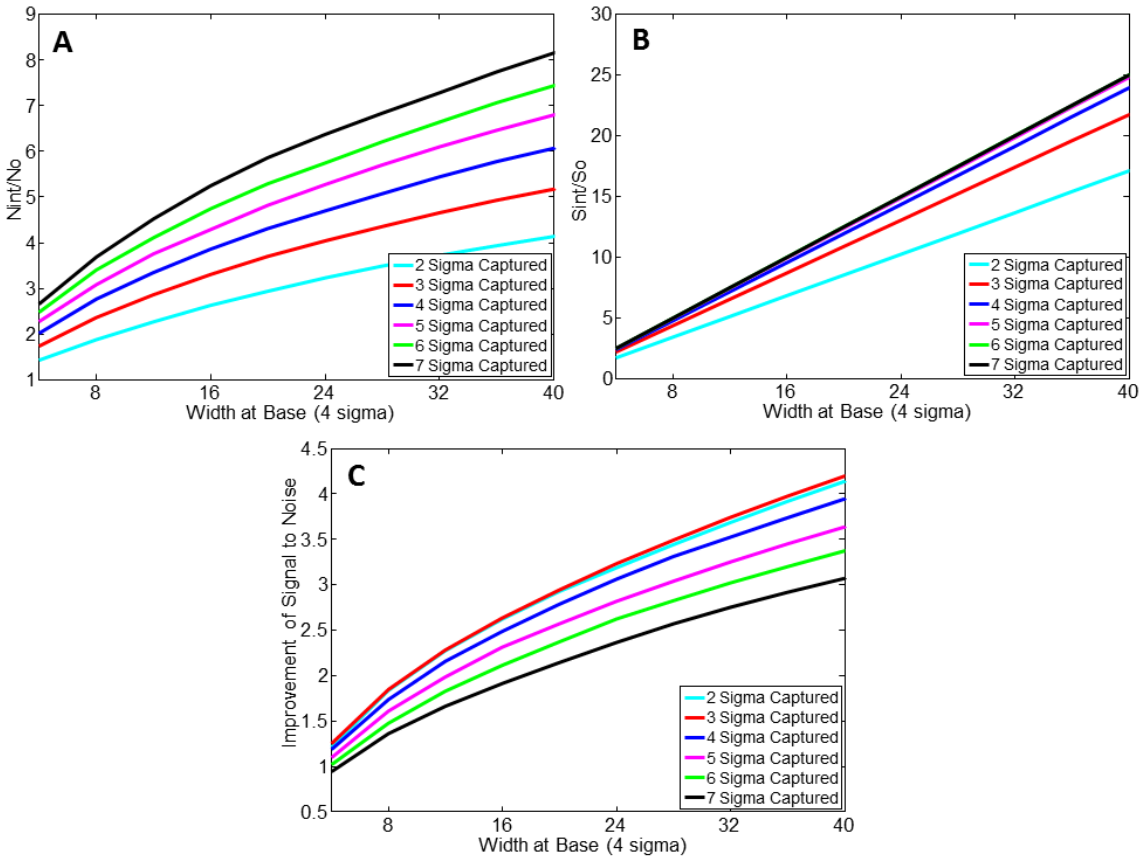


Figure 10.3 (A) Random noise was generated and for various peak widths (4σ), different amount of variance was captured by integrating across various amounts of σ. Herein the noise has been divided by the traditional definition of noise to show the increase in the noise. (B) For integrating a peak, a larger amount of signal will be captured if your bounds are larger (i.e. if you integrate with a sigma captured of 7, 99.7% of the peak is captured while integrating capturing two sigma will only capture 68.2% of the peak). Based on equation, $\frac{S_{int}}{S_0} = \sigma\sqrt{2\pi^{1/2}}$ the integration signal increases linearly with peak width (C) The signal improvement curves (B) have been divided by the increase in the noise curves (A) yielding a improvement in the signal to noise as a function of peak width. Based on the curves, capturing 3σ will yield an optimum improvement in the S/N. Commonly peaks are roughly 20 data points wide at the base which when integrated will have a S/N improvement of a factor of 3.

The LOD which is defined as the height of the peak three times the standard deviation of the noise in traditional terms. Shown in Figure 10.4A are ten curves which are a subset of 1000 curves that were generated with an average *S/N* of 3. By integrating the peaks and the *S/N* has been increased to 7.2 when 6σ is captured or 9.0 when 3σ is captured as shown in Figure 10.4B. By extending the LOD by a factor 3, analytes can be detected at lower concentrations. Furthermore the LOQ is also extended by a factor of 3 allowing more sensitive and accurate quantitation of analytes at low concentrations. We argue that to determine if a peak is present, one should integrate and capture 4σ of the peak. However, one should capture 6σ of the peak when determining the concentration of the peak as this captures more of the peak area and is a more accurate quantitation of the peak.

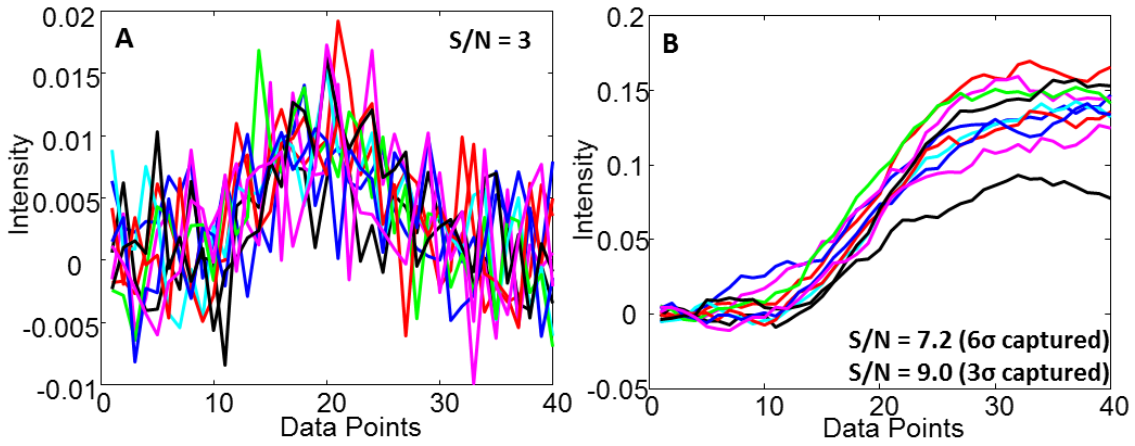


Figure 10.4 (A) A chromatographic peak has been modeled by Gaussian distribution which has a W_b of 20 data points with a S/N of 3. This has been traditionally defined as the limit of detection. Ten replicates are shown (B) The peaks have been integrated, and the S/N has been increased 2.4 times for an 6σ captured or a 3.0 times increase in S/N for 3σ sigma captured. The peak is now significantly above the LOD.

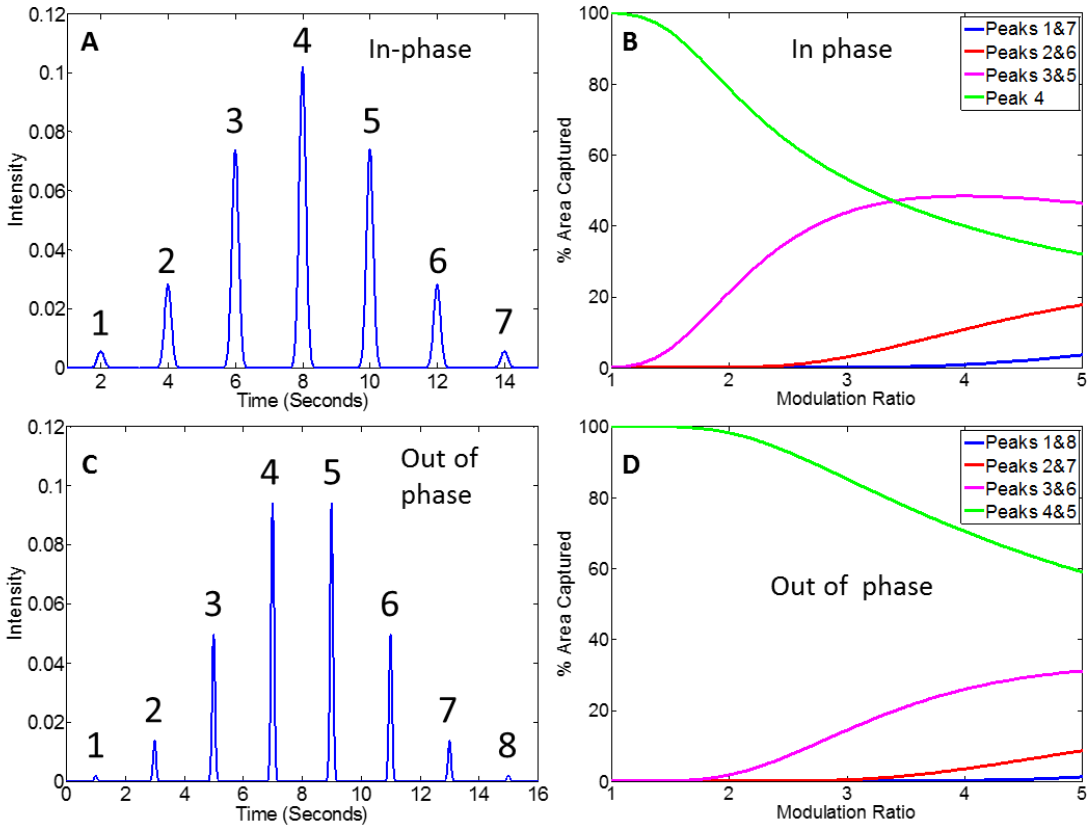


Figure 10.5 (A) A two dimensional separation is shown with a M_R of 5 for an in-phase peak. The modulations are labeled (B) The amount of area contained as a percentage of the total area per modulation of the original peak is shown for various M_R. For a M_R of one, Peaks 1,2,3,5,6, and 7 are infinitesimally small meaning almost 100% of the area is contained within Peak 4. As the M_R increases, the area of the original ^{1}D peak area is spread out more amongst the modulations. (C) A two dimensional separation is shown with a M_R of 5 for an out of phase peak. The modulations are labeled (D) Amount of area contained as a percentage of the total area per modulation of the original peak is shown for various M_R for an out of phase peak. Similar to (B), at lower M_R most of the area is contained within peaks 4 and 5 while at higher M_R the area of the ^{1}D peak is spread out to other modulations.

This integration method can also be applied to two-dimensional separations where it can be useful to determine which peaks on the ^{2}D to include or exclude when determining concentration of analytes. As the modulation ratio (M_R) increases, the area under the one-dimensional peak is spread out between more of the ^{2}D peaks. At a M_R of one for an in-phase peak, almost 100% of the peak area is contained within one peak. However, at a M_R of 5, the middle ^{2}D peak only contains 46 % of the area. Leaving 54 % of the area to be spread out amongst 5 other ^{2}D peaks. Often the outer ^{2}D peaks which contain only 3.58 % of the area are excluded due to being below the LOQ which can lead to inaccurate quantitation for the analyte. For an out of phase peak, more of the area is contained within inner two modulations which means that not including the area from the smaller ^{2}D peaks could be more accurate compared to in-phase peaks. Nevertheless, not including the outer ^{2}D peaks will results in less accurate quantitation of the analyte. A summary of how M_R affects the distribution of the peak area for both in-phase and out of phase peak can be seen in Figure 10.5.

To apply the integration method to two-dimensional data, the peak (or peaks) in question should be extracted from the chromatogram and converted into a one-dimensional view as seen in Figure 10.6A. One can always select a larger area and more modulations to ensure that the entire peak is captured. Because the data was modeled, we know the exact W_b (20 data points) and location of ^{2}D peak maxima. As seen in Figure 10.6B, the extracted information from the two dimensional data has been converted into one dimensional data and 8σ (40 data points) is included for each peak. Figure 10.6C shows the peaks after integration. Modulations 2 and 6 where originally below the LOD but now are significantly above the LOD after integration. After integration modulations 3 and 5 are above the LOQ and can be included in the calculation of the

peak area for purposes of finding analyte concentration. See supporting information (Fig E.1) for the individual integration curves that are excluded for brevity.

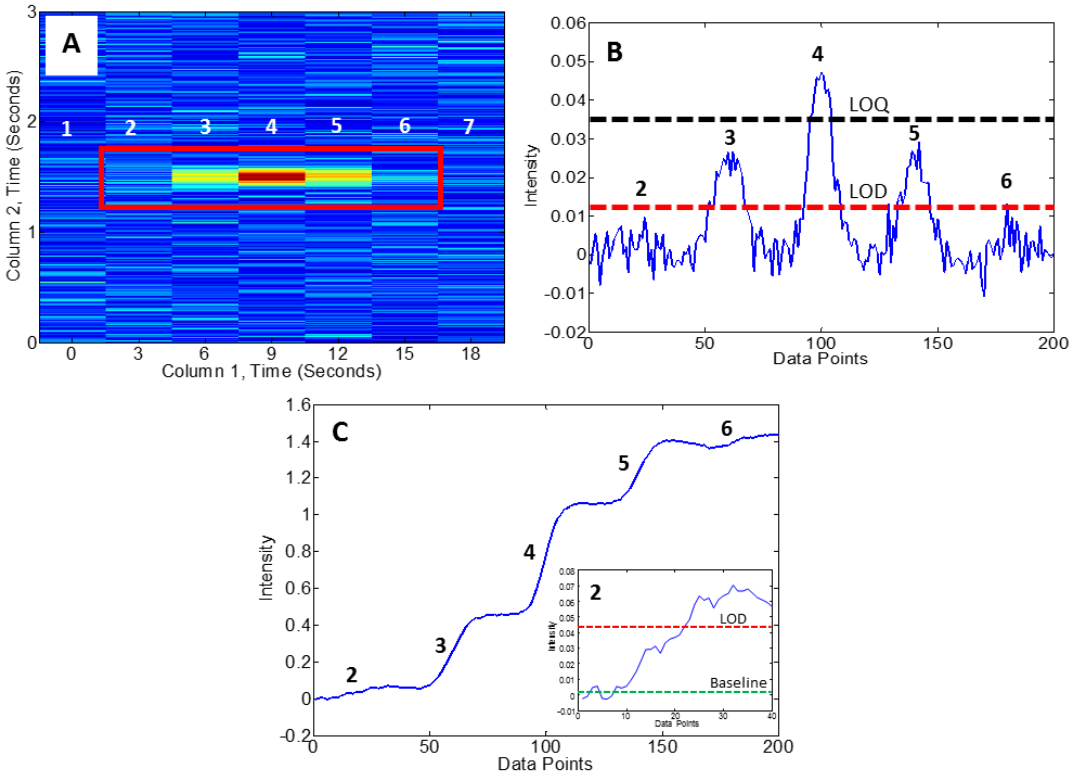


Figure 10.6 (A) modeled two dimensional chromatographic peak is shown with M_R of 4. The area within the box was cut out. (B) The area that was cut out of the box was converted to a one dimensional view. The peaks have widths of 20 data points. and the most intense peak has a S/N of 14 (C) The data was integrated and the S/N of each individual peak has been increased roughly 3 times. Thus peak 2 is now above the LOD as seen in the insert as well as peak 6. Peaks 3 and 6 are now significantly above the LOQ (see Figure E.1, Supporting Information). Peak 4 was barely above the LOQ in original data but now is significantly above the LOQ.

Concentrations of analytes are often determined using the two step method. Briefly, this method entails looking at the maximum intensity of the peak of interest. If this intensity is above the LOQ, the peak is integrated. If the signal of the peak is not high enough it is excluded from determining the concentration of the analyte, but can used to identify the analyte if above the LOD. In two dimensional separations, this often means that modulations are excluded when determining the concentration of the analyte leading to inaccurate quantitation. Traditionally, peaks have only been included in the quantitation of an analyte if the signal is above the LOQ. At this *S/N*, there is an extremely small chance that the measured signal will fall within the distribution of the noise. Conversely at the LOD, there is a 0.13% chance that the measured signal will fall within the distribution of the noise. An illustration is shown in Figure 10.7 comparing using the two step method and the integration method using LOD and LOQ to quantify the concentration of an analyte. At high *S/N* (>100), both methods, two step and integration, are acceptable to use. Both methods yield accurate and precise quantitation. At moderate *S/N* (≈ 70) of the most intense modulation measured in the traditional means, the two step method starts to exclude some of the smaller peaks. At a M_R of 4 for an in phase peak (as seen in Figure 10.7), the outer ^{2}D peak (peaks 2 and 6 as seen in Figure 10.5) account for roughly 11% of the signal. The two step-method excludes these peaks when the most intense modulation is roughly at a *S/N* of 80 while the integration method using the LOQ starts to exclude these peaks at roughly a *S/N* of 45. At a *S/N* of roughly 20, the two step method starts to exclude peaks 3 and 5 which results in about a 60% error while the integration method using LOQ starts to fail when the *S/N* of the most intense modulation is roughly 10. It is apparent that the integration method is more accurate than the two step method at lower concentrations. However, if we use the integration method and quantify all peaks if they are above the LOD, a more accurate answer is achieved. For a M_R of 4 for an out of

phase peak, the integration method using LOQ captures nearly 100% of the peak area to about a *S/N* of 17 for the most intense modulation. At lower *S/N*s the integration method using LOQ is inaccurate, but quantifying anything above the LOD we achieve an answer closer to the true area. The precision is lower but not significantly so except for at extremely low *S/N*s (*S/N* ≈ 3). We would argue that having a more accurate answer with a small loss in precision is better than excluding peaks by either method, two step or integration using LOQ. Further comparison of the two step method to the integration method at various M_R can be seen in the supporting information (Figure E.2).

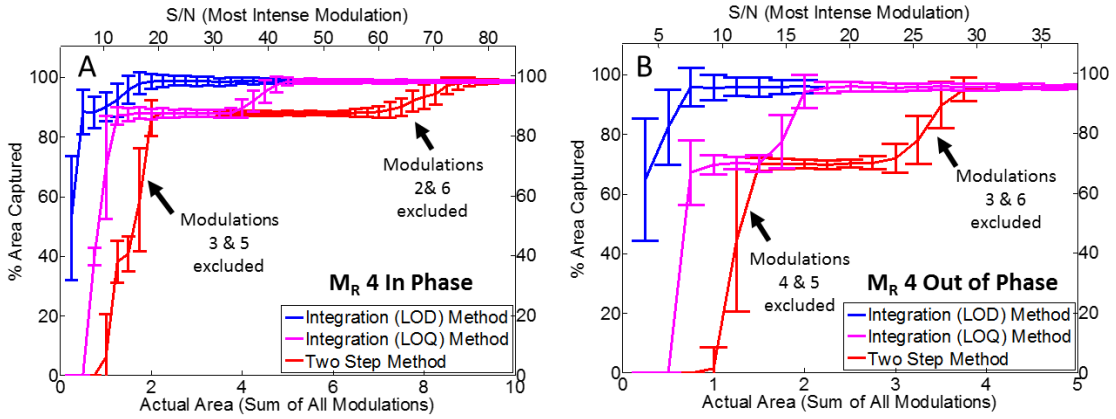


Figure 10.7 (A) Comparison of the integration method and two step method for a M_R of 4 for an in phase peak. At high S/N, both techniques work quite well capturing almost 100% of the area with similar errors. Modulations 1 & 7 contain about 1% of the total peak area. When modulation 4 reaches a signal to noise around 80, modulations 2 & 6 start to be excluded because they are not above the LOQ. Because the integration method has an increase in S/N modulations 2 & 6 will be included at a lower S/N. At a S/N of ~35 the integration method will exclude the 2nd and 6th modulation as well. At a S/N of ~18 the two step method will start to exclude modulations 3 & 5. Quantifying peaks above the integration LOD gives the most accurate answer but is not as precise at lower S/N (B) Shown is the comparison of the three methods for an out of phase peak at an M_R of 4. Similar results are observed as (A).

In order to demonstrate this technique on non-simulated data, a test mixture was analyzed via GC × GC – TOFMS. Figure 10.8A shows the chromatogram for bromobenzene at *m/z* 156. For quantification of mass spectral data, a specific *m/z* is used which is at high *S/N* and is not interfered. The portion of the chromatogram within box in Fig. 10.8A is extracted and has been converted into the raw vector form in Fig 10.8B where Savitzky-Golay smoothing was performed. The noise of the raw vector was determined to be 4.47 and the $^{2}W_{b}$ is 70 data points (70 ms). There are four modulations which are above the LOQ; however, there is a fifth modulation centered at 235 data

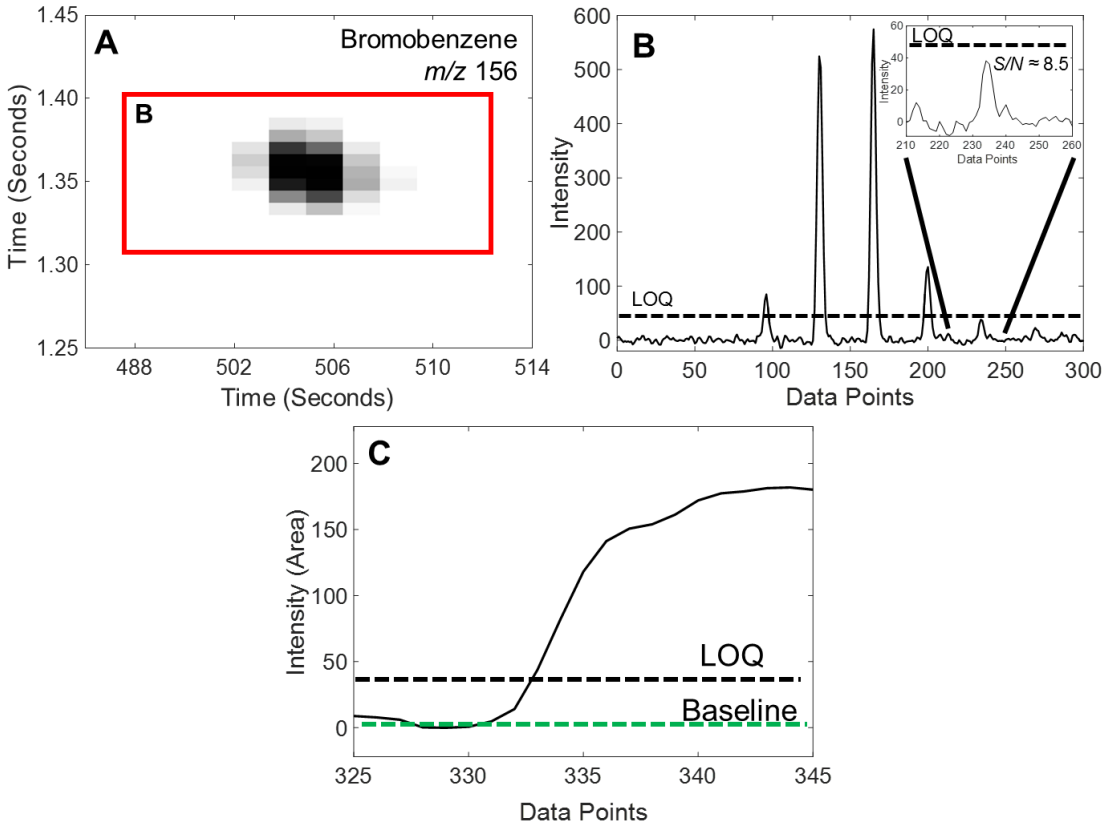


Figure 10.8 Application of the integration method to non-simulated data. (A) GC × GC chromatogram of bromobenzene using *m/z* 156. (B) Vector form for bromobenzene which has been extracted from the GC × GC chromatogram. The four most intense modulations are above the LOQ but the fifth is slightly below. (C) Integration of the least intense modulation which is now significantly above the LOQ.

points which is above the LOD and below the LOQ with a $S/N \approx 8.5$. This peak was integrated as shown in Fig 10.8C. The noise in the integrated domain was determined to be 4.07 which is significantly lower than expected. Thus the *S/N* in the integrated domain was determined to be 40 which is 4 fold increase from the traditional measurement of *S/N*.

10.5 CONCLUSION

In this report, we have provided a new, novel way to analyze one and two dimensional chromatographic data. By using the proposed integration method, sensitivity can be increased by a factor of 3. This increased sensitivity allows for quantitation of peaks at lower S/N compared to other methods of analysis. We have presented a thorough investigation of the manifestation of the noise in the integrated domain and how to achieve the maximum improvement in S/N. Furthermore, we have shown, that the integration method is a more accurate method to quantitate two-dimensional data. Finally, we investigated using this method for both LOD and LOQ. Using the LOD to determine which peaks to include results in more accurate answers with only a minor loss to the precision. Further investigation of this method involved analysis of GC × GC – TOFMS data for practical analysis. This integration method can be easily incorporated into typical data handling systems for chromatographic data. Finally, this method has general applicability to any data set which has a Gaussian distribution as a function of time and uncorrelated, normal distribution of noise.

10.6 SUPPORTING INFORMATION

Figures E.1 and E.2 can be found in Appendix E

10.7 REFERENCES

[1] H. Kaiser, Quantitation in elemental analysis, Anal. Chem. 42 (1970) 24A–41A.
[2] P.W.J.M. Boumans, A tutorial review of some elementary concepts in the statistical evaluation of trace element measurements, Spectrochimica Acta Part B: Atomic Spectroscopy. 33 (1978) 625–634.

[3] L.A. Currie, The limit of precision in nuclear and analytical chemistry, Nuclear Instruments and Methods. 100 (1972) 387–395.
[4] E. Bernal, Limit of Detection and Limit of Quantification Determination in Gas Chromatography, (2014).
[5] J. Krupčík, P. Májek, R. Gorovenko, J. Blaško, R. Kubinec, P. Sandra, Considerations on the determination of the limit of detection and the limit of quantification in one-dimensional and comprehensive two-dimensional gas chromatography, Journal of Chromatography A. 1396 (2015) 117–130.
[6] K.H. Chang, Limit of Detection and Its Establishment in Analytical Chemistry, Health and the Environmental Journal. 2 (2011) 38–43.
[7] Nomenclature, Symbols, Units and Their Usage in Spectrochemical Analysis-II. Data Interpretation, International Union of Pure and Applied Chemistry, Spectrochimica Acta Part B. 33 (1978) 241–245.
[8] Y.V. Heyden, R. Boque, Practical Data Handling-The Limit of Detection-The concept of the limit of detection (LOD) often arouses controversy in analytical chemistry, LCGC Europe. 23 (2009) 82–85.
[9] J.E. Knoll, Estimation of the Limit of Detection in Chromatography, J Chromatogr Sci. 23 (1985) 422–425.
[10] D.G. Peters, J.M. Hayes, G.M. Hieftje, Chemical separations and measurements: theory and practice of analytical chemistry, WB Saunders Company, 1974.
[11] J.D. Wineforder, Trace Analysis: Spectroscopic Methods for Elements, Wiley, New York, 1976.
[12] G.L. Long, J.D. Wineforder, Limit of Detection: A Closer Look at the IUPAC Definition, Analytical Chemistry. 55 (1983) 712A–724A.
[13] D.A. Armbruster, M.D. Tillman, L.M. Hubbs, Limit of detection (LQD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused drugs., Clinical Chemistry. 40 (1994) 1233–1238.
[14] J.P. Foley, J.G. Dorsey, Clarification of the limit of detection in chromatography, Chromatographia. 18 (1984) 503–511.
[15] W. Khummueng, J. Harynuk, P.J. Marriott, Modulation Ratio in Comprehensive Two-dimensional Gas Chromatography, Anal. Chem. 78 (2006) 4578–4587.
[16] R.E. Synovec, E.S. Yeung, Improvement of the limit of detection in chromatography by an integration method, Anal. Chem. 57 (1985) 2162–2167.
[17] E.S. Page, Continuous Inspection Schemes, Biometrika. 41 (1954) 100–115.
[18] E.S. Page, Cumulative Sum Charts, Technometrics. 3 (1961) 1–9.
[19] N.L. Johnson, A Simple Theoretical Approach to Cumulative Sum Control Charts, Journal of the American Statistical Association. 56 (1961) 835–840.
[20] R.B. Crosier, A New Two-Sided Cumulative Sum Quality Control Scheme, Technometrics. 28 (1986) 187–194.
[21] J.M. Lucas, R.B. Crosier, Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM A Head Start, Technometrics. 24 (1982) 199–205.
[22] D.M. Hawkins, D.H. Olwell, Cumulative Sum Charts and Charting for Quality Improvement, Springer Science & Business Media, 2012.
[23] A.B.J. Nix, S.J. Rowland, K.W. Kemp, D.W. Wilson, K. Griffiths, Internal quality control in clinical chemistry: A teaching review, 6 (1987) 425–440.
[24] H.-H. Kuo, White Noise Distribution Theory, CRC Press, 1996.

[25] J. Poloniecki, O. Valencia, P. Littlejohns, Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery, BMJ. 316 (1998) 1697–1700.
[26] P.H. Ellaway, Cumulative sum technique and its application to the analysis of peristimulus time histograms, Electroencephalography and Clinical Neurophysiology. 45 (1978) 302–304.
[27] R.A.G. Marshall, Cumulative sum charts for monitoring of radioactivity background count rates, Anal. Chem. 49 (1977) 2193–2196.
[28] H.S. Johnston, C.E. Miller, B.Y. Oh, K.O. Patten, W.N. Sisk, Internal energy distributions from nitrogen dioxide fluorescence. 1. Cumulative sum method, J. Phys. Chem. 97 (1993) 9890–9903.
[29] A. Nijhuis, S. de Jong, B.G.M. Vandeginste, The application of multivariate quality control in gas chromatography, Chemometrics and Intelligent Laboratory Systems. 47 (1999) 107–125.
[30] E. Rubi, R.A. Lorenzo, M.C. Casais, R. Cela, Quality control in the routine analysis of methylmercury in biological and environmental materials using gas chromatography with electron capture detection, Applied Organometallic Chemistry. 8 (1994) 677–686.

Chapter 11. Conclusion

11.1 SUMMARY OF PRESENTED WORK

The evolution from one-dimensional gas chromatography (GC) to comprehensive multi-dimensional gas chromatography (MDGC), namely two-dimensional gas chromatography (GC × GC), has been one of the most significant advancements in the field of GC in the past 27 years. In addition to the improvements in instrumentation, there have been numerous developments to optimization, materials and methodology, and data processing and analysis. A multitude of researchers have applied numerous approaches to improve the various components that make up a GC × GC separation and data analysis techniques, resulting in significant improvements. The presented work developed new instrumentation and data analysis techniques that further expand the chemical information that can be gleaned from MDGC with a focus on GC × GC.

Chapters 2 through 4 focus on new modulators for GC × GC that are robust, reliable, cost-effective, and still maintain similar figures of merit (peak capacity and peak capacity production) to commercially available systems. Through use of a pulse flow valve, an unconventional form of modulation, partial modulation, was investigated that provided a peak capacity on the ^{2}D dimensional ($^2n_c = 10$) while achieving this feat with a modulation period (P_M) of only 250 ms. Furthermore, the pulse flow valve was shown to operate with P_M of 50 ms enabling it to be used for high-speed separations as well as used as a "GC sensor." While this study was very preliminary, further research with this pulse flow valve is warranted. Chapters 3 and 4 investigated high temperature diaphragm valve-based modulators for GC × GC – FID and GC × GC – TOFMS. Separation temperatures up to 325 °C were achieved, and narrow, reproducible, and symmetric peaks were created on the ^{2}D dimension. Peak capacity of 9600 in ~30 minutes was achieved for

GC × GC – FID separations, and a peak capacity of ~1,500 in ~20 minutes was achieved for the GC × GC –TOFMS separations. The knowledge gained from the investigation of the pulse flow valve and high temperature diaphragm valve were applied to GC^3 as demonstrated in Chapter 5. Building on previous work by the Synovec group, an optimized instrumental platform was designed allowing for both increased selectivity provided by the three dimensions and a high peak capacity production of 1,000 peaks/min, a 4-fold improvement from previous GC^3 platforms.

Chapter 6 focused on the implementation of a cost-effective instrument for routine analysis of kerosene-based rocket fuels. A high temperature diaphragm valve was used as a modulator for GC × GC – FID, and partial least squares (PLS) analysis was used to correlate the chemical information of fuels obtained via GC × GC – FID to several chemical and physical properties of the fuels. The models had low RMSECV values and interpretation of the linear regression vectors (LRVs) allowed for identification of compounds that positively or negatively correlated with the properties of interest.

Chapter 7 and 8 also applied PLS to link chemical information obtained via GC × GC – TOFMS to physical properties of interest. Chapter 7 employed an extremely large, diverse set of kerosene-based fuels (74 samples) in order to gain a better fundamental understanding of the relationship between chemical composition and fuel properties such as kinematic viscosity, heat of combustion, hydrogen content, and density at three different temperatures (15 °C, 45°C, and 85 °C). The PLS modeling yielded highly reliable models with errors ranging from ~6% to 13%. Investigation of the linear regression vectors (LRVs) indicated the relationship between the chemical composition and physical properties enabling the chemical compositions of fuels to in principle be altered to meet desired application specifications. Chapter 8 used a subset of these fuels (36 samples) and investigated the relationship between chemical composition and thermal

stability. The thermal stability of the fuels was evaluated by CRAFTI and subsequent carbon deposition analysis. Initial modeling using the entire chromatograms did not produce satisfactory models. Thus, a novel feature selection technique was implemented resulting in improved models as well as a better understanding of the relationship between the chemical composition and thermal stability.

The complementary information provided by tandem ionization time-of-flight mass spectrometry (TI – TOFMS) was leveraged to improve the discovery-based analysis in Chapter 9. Twelve analytes that were spiked into diesel fuel were "discovered" using the tile-based Fisher ratio software. The hard and soft ionization energies (70 and 14 eV) hit lists and the fusion of the two datasets were investigated resulting in improved discovery. Eleven of the twelve analytes were "discovered" and using PARAFAC deconvolution, nine of the twelve analytes were identified using spectral matching. The limit of discovery and limit of identification were investigated and it was concluded that tandem ionization lowered these thresholds.

In the Chapter 10, a novel method for lowering the limit of detection and quantification (LOD and LOQ) were explored for one- and two-dimensional chromatography. Using the novel integration method, the LOD and LOQ were lowered by a factor of 3 for one-dimensional chromatography. This was then applied to two-dimensional chromatography and was compared to the traditionally used two-step method. Application of this technique to a "real-world" sample collected via two-dimensional gas chromatography with a time-of-flight mass spectrometer (GC × GC – TOFMS) resulted in a *S/N* improvement of 36-fold which exceeded the theoretical predications.

11.2 FUTURE DIRECTIONS

There remain several future directions for the research presented in each chapter of this document. This section will expand on some potential ideas for new research.

11.2.1 *Future Directions for Instrumentation*

There is a large untapped reservoir of research for the pulse flow valve as a modulator. Indeed, several ideas have already been proposed for subsequent graduate students to pursue at the time of the writing of this manuscript. The first would be using the pulse valve to continuously monitor a reaction in real time. A continuous stream of gas from the headspace of the reaction could be piped into the GC and the pulse valve could be actuated at a relatively rapid rate (<75 ms). The resulting chromatogram could be "cut" at each modulation resulting in a series of 1D separations that are representative of the reaction. PCA could be used to simplify the analysis of the data, and the resulting scores plot should be a function of the time points. The loadings plots would provide chemical insight into the reaction profile as well as further insight into reaction mechanisms and reaction rates.

While the original intention of the pulse flow valve was to find an ultrafast modulator which would enable high speed separations, a downside to its use is the requirement to differentiate the data in order to visual the chemical information. Current research being performed by H.D. Bahaghighat has shown that it is possible to use multivariate curve resolution alternating least squares (MCR-ALS) to deconvolute GC × GC – TOFMS in order to extract the chemical information without processing the data. Nevertheless, there continues to be numerous opportunities to create high speed modulators that perform MDGC in the traditional approach by creating Gaussian peaks or by introducing disturbances into the ^{1}D Gaussian profiles creating "sensor" information.

Research with the high temperature diaphragm valves presented herein and research performed by others (Watson et. al) have shown the diaphragm valve to be a modulator system that is robust, reliable, and cost-effective while still yielding identical performance to other state-of-the-art systems as well as commercially available systems. Many researchers across the world are recognizing that there are alternative modulation systems compared to the traditional liquid nitrogen (LN_2) systems. Diaphragm valve technology is ready to be commercialized and introduced to the wider GC × GC community.

11.2.2 *Future Directions for Data Analysis*

Partial least squares (PLS) analysis has been shown to be an effective tool in relating chemical information obtained from GC × GC – FID and GC × GC – TOFMS to physical properties of interest. All the PLS analysis herein was performed on kerosene-based aeronautical fuels, but this technique does not need to be limited to such a specific sector of research. Expansion of this technique to food chemistry, metabolomics, and many other areas would provide valuable insight into the relationship between chemical information and physical properties. PLS can also be used outside the field of GC × GC. As seen in Chapter 8, PLS may not always work unless a feature selection technique is implemented. The use of feature selection for PLS has been rather limited and there is only one publication for feature selection for PLS in GC × GC. Chapter 8 introduces a novel algorithm for feature selection which could be expanded to other modeling situations when PLS does not work initially or the resulting models do not yield favorable results. Tandem ionization (TI) is in its early stages as a detection technique, and there are numerous areas to investigate. Deconvolution with TI needs to be further investigated as having two “sets” of data may be beneficial in lowering the required resolution to deconvolute severely overlapped peaks. Very preliminary research has been performed about the complementary information provided by

TI. Match values and spectral contrast angles were calculated between the low and high ionization energies (14 and 70 eV) for a diesel fuel sample as shown in Figure 11.1. The alkane species have a high match value and a low spectral contrast angle. The cycloalkanes have a relatively high match value and a moderate spectral contrast angle, and the aromatics have a lower match value and a high spectral contrast angle. This trend might be present because alkanes fragment a large amount at both 70 and 14 eV and thus have relatively similar spectra. Aromatics are a more stable entity and thus the higher ionization voltage will result in more fragmentation while the low ionization voltage will result in minimal or almost no fragmentation. Thus, there will be a lower match value and higher spectral contrast angle. Further investigation into this pattern is warranted.

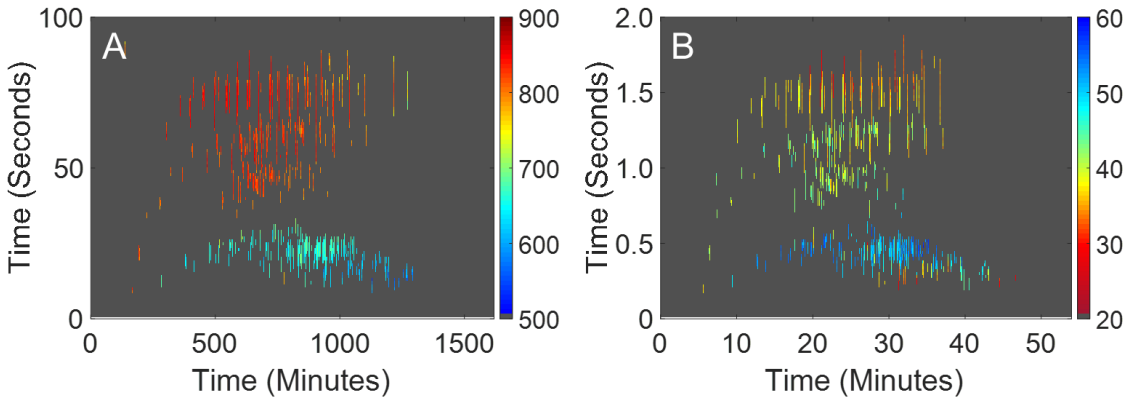


Figure 11.1 (A) The match value comparing the 70 to 14 eV mass spectrum was calculated at every pixel above a TIC signal threshold of 40,000. The alkanes and cycloalkanes mass spectra are similar to each other while the aromatics do not have very high match values. (B) The spectral contrast angle comparing the 70 to 14 eV mass spectrum was calculated at every pixel above a TIC signal threshold of 40,000. The same trend for the alkanes and aromatics is seen using the spectral contrast angle.

Appendix A

Table A.1. Table of the 15-component mixture ordered by elution time on ^{1}D. The supplier and purity are included. The 15 compounds were combined equally by volume by adding 25 µl each for a total volume of 375 µl.

Elution	Name	Boiling Point (°C)	Supplier	Purity (%)
1	Acetone	56	Sigma-Aldrich	99.9
2	Isopropyl Alcohol	83	Sigma-Aldrich	99.7
3	2-Methyl-2-propanol	82	Sigma-Aldrich	99.5
4	Cyclopentane	49	Sigma-Aldrich	99.0
5	1-Hexene	63	Sigma-Aldrich	97.0
6	2-Butanone	80	Sigma-Aldrich	99.0
7	Hexane	68	Sigma-Aldrich	97.0
8	1-Hexyne	71	Sigma-Aldrich	97.0
9	Methylcyclopentane	72	Honeywell	97.0
10	1-Chlorobutane	79	Sigma-Aldrich	99.5
11	Carbon Tetrachloride	77	Sigma-Aldrich	99.9
12	Benzene	80	Fisher Chemical	99.0
13	Cyclohexane	81	Sigma-Aldrich	99.0
14	Heptane	98	Fisher Chemical	99.4
15	1-Heptene	94	Sigma-Aldrich	97.0

Table A.2. Table of the 18-component mixture ordered by elution time on ^{1}D. The supplier and purity are included. The 18 compounds were combined equally by volume by adding 25 µl each for a total volume of 450 µl.

Elution	Name	Boiling Point (°C)	Supplier	Purity (%)
1	Isopropyl Alcohol	83	Sigma-Aldrich	99.7
2	Acetone	56	Sigma-Aldrich	99.9
3	Hexane	68	Sigma-Aldrich	97.0
4	1-Hexene	63	Sigma-Aldrich	97.0
5	t-butyl methyl ether	55	Sigma-Aldrich	99.8
6	2-Methyl-2-propanol	82	Sigma-Aldrich	99.5
7	Cyclopentane	49	Sigma-Aldrich	99.0
8	Methylcyclopentane	72	Honeywell	97.0
9	1-Hexyne	71	Sigma-Aldrich	97.0
10	Carbon Tetrachloride	77	Sigma-Aldrich	99.9
11	Heptane	98	Fisher Chemical	99.4
12	Ethyl acetate	77	J.T.Baker	99.9
13	Cyclohexane	81	Sigma-Aldrich	99.0
14	1-Heptene	94	Sigma-Aldrich	97.0
15	2-Butanone	80	Sigma-Aldrich	99.0
16	1-Chlorobutane	79	Sigma-Aldrich	99.5
17	Benzene	80	Fisher Chemical	99.0
18	1,2-Dichloroethane	83	Sigma-Aldrich	99.8

Table A.3. List of 8 compounds that were spiked into diesel fuel in order of elution time on ^{1}D. 100 µl of each compound was mixed to create a 800 µl stock solution. Then 10 µl of the stock solution was added to 190 µl of diesel fuel to create a 200 µl solution.

Elution	Name	Boiling Point (°C)	Supplier	Purity (%)
1	1,1,1-Trichloroethane	74	Sigma-Aldrich	99.5
2	Butylamine	78	Sigma-Aldrich	99.8
3	1,2-Dichloroethane	83	Sigma-Aldrich	99.8
4	1,6-Dichlorohexane	90	Sigma-Aldrich	98.0
5	N-Cyclohexylaniline	192	Sigma-Aldrich	99.0
6	N,N-Dimethylaniline	194	Sigma-Aldrich	99.0
7	m-Toluidine	203	Sigma-Aldrich	99.0
8	Dibenzylamine	300	Sigma-Aldrich	97.0

Figure A.1. Modulation of ^{1}D signal with pulse flow valve and processing steps (3 steps) to convert signal to ^{2}D format with the appearance of a GC × GC chromatogram. (A) The unmodulated peak of a non-overlapped single analyte. (B) The same analyte after partial modulation with the pulse flow valve. (C) Differentiation of the partially modulated peak in (B). The minima represents the inflection point of the error function produced by the pulse valve flow modulator. (D) Inversion of the differentiated data. This is performed by multiplying the differentiated by -1. (E) The baseline sag produced by the differentiation step is removed by an in-house low frequency noise filter. (F) A 2D chromatogram produced by cutting and folding the data by the modulation period (P_M) time.

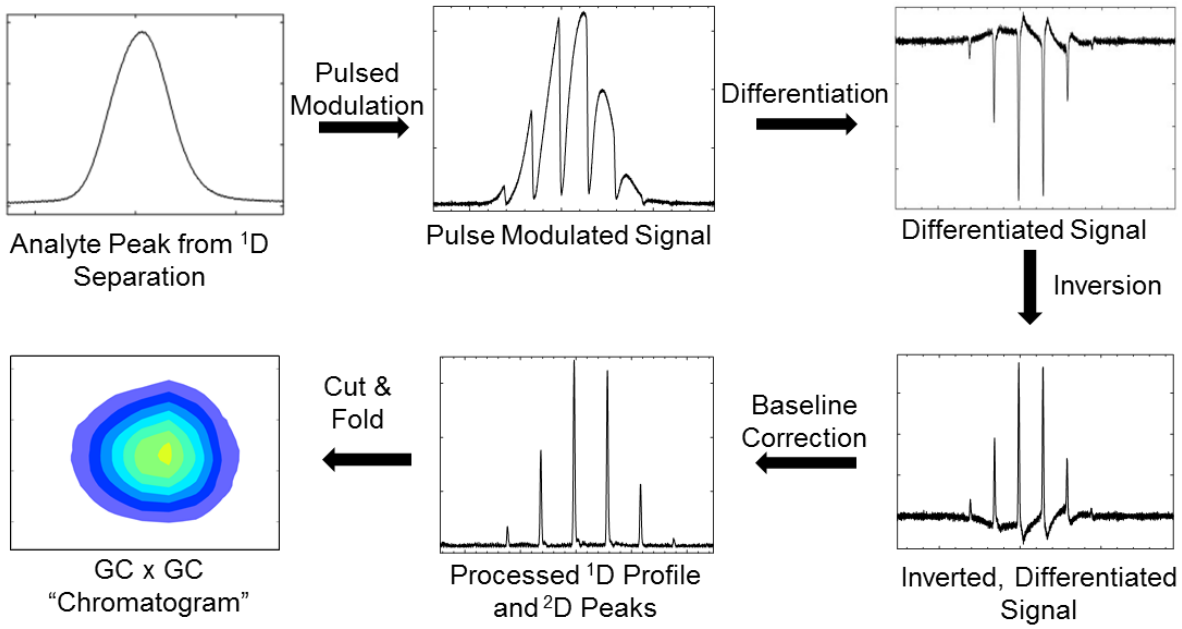


Figure A.2. (A) Shown is a 2D Chromatogram of the 115 mixture separation shown in Figure 5.2. Twelve analytes were used to evaluate the peak capacity on the second dimension (^2n_c). (B) The twelve analytes were cut out of the 2D chromatogram, concatenated, and are plotted in a single chromatogram to measure their respective ^{2}D retention time and peak width (^2W_b). Retention times verses peak width is shown in the subset along with the fitted equation. (C) A model based on the equation from (B) is used to calculate the ^2n_c. A resulting second dimension peak capacity of 20 was calculated at a resolution of 1, for a total peak capacity $n_{c,2D}$ of ~7200 or a peak capacity production of ~1200 peaks/min for this separation.

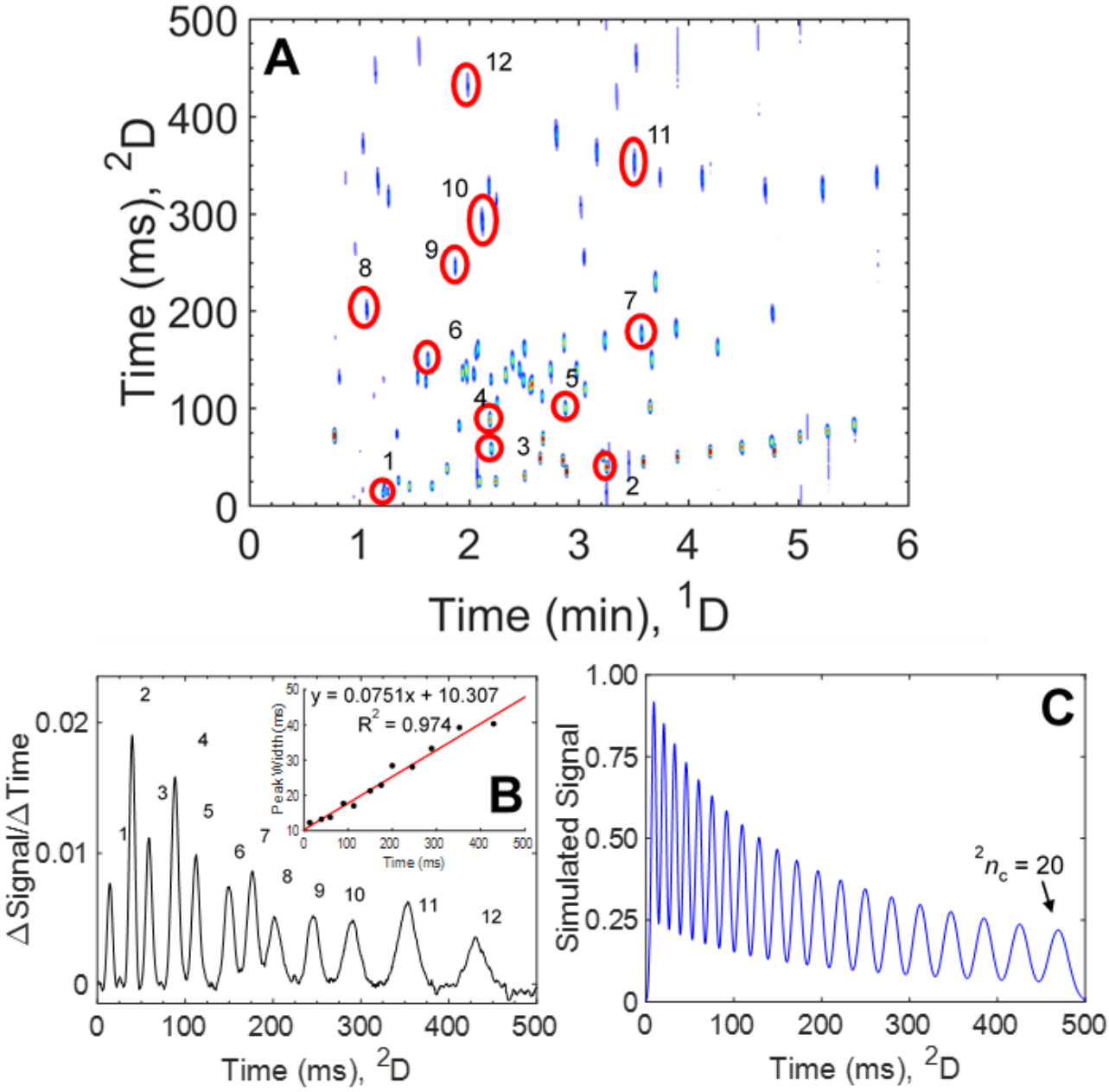


Figure A.3. (A) The combined processed vector chromatogram of a GC^3 - FID separation of the 18 component mixture. Each analyte was injected individually to identify the analyte based on retention times. The ${}^1P_M$ was 1 sec and the ${}^2P_M$ was 50 ms. (**B**) An enhanced view of the processed vector chromatogram of is shown for n-Hexane. The width of based data is listed for all three separation dimensions. (**C**) An enhanced view of the processed vector chromatogram of is shown for Methylcyclopentane (red), and 1-Hexyne (yellow). The width of based data is listed for all three separation dimensions.

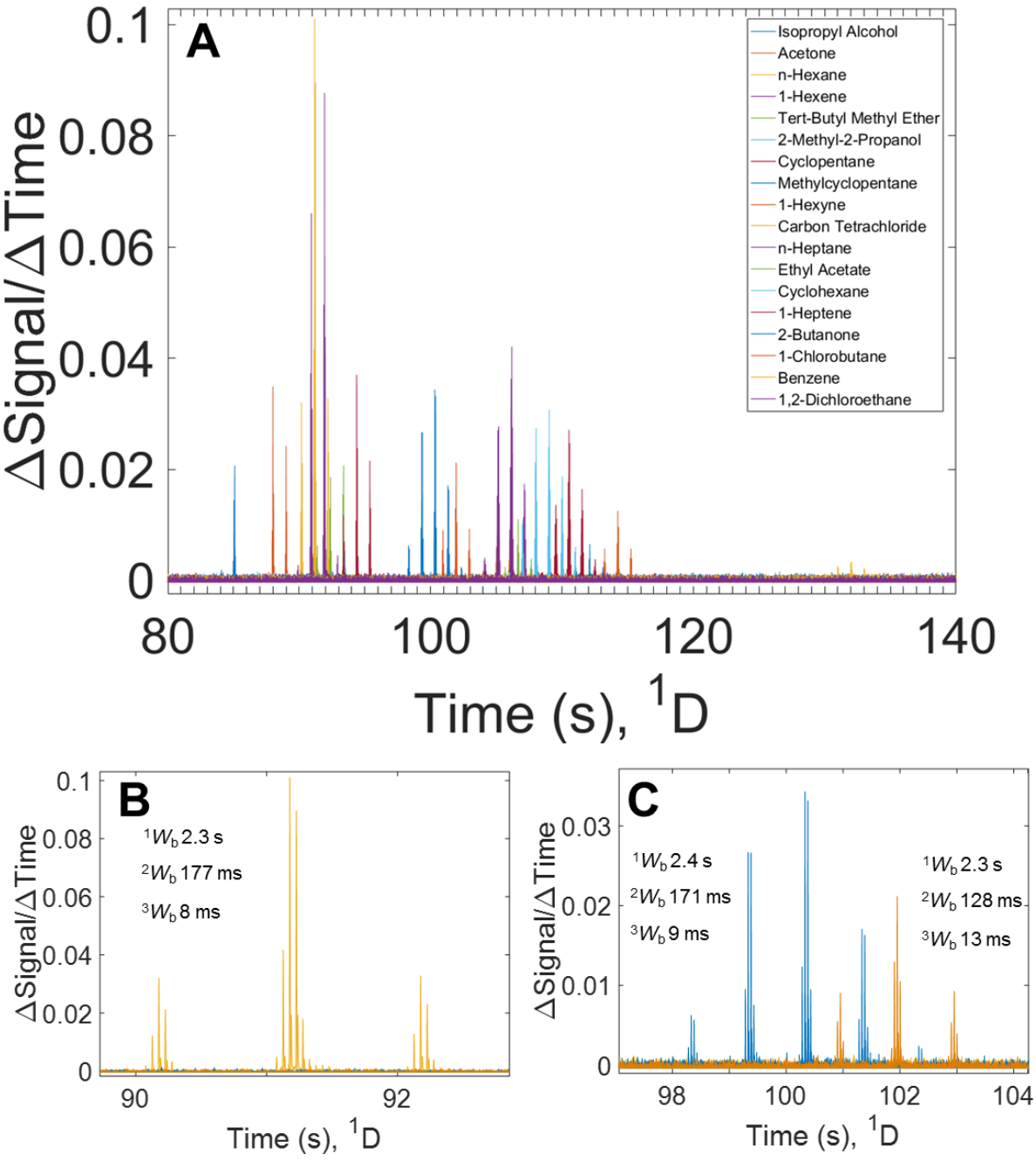


Figure A.4. (A) ^{1}D vs ^{2}D chromatogram of the 18 component test mixture collected using a $^{1}P_{M}$ of 1.0 seconds and a $^{2}P_{M}$ of 50 ms that has been summed on the ^{3}D. The analytes from Table 5.1 have been circled. (B) ^{1}D vs ^{3}D chromatogram of the 18 component test mixture collected using a $^{1}P_{M}$ of 1.0 seconds and a $^{2}P_{M}$ of 50 ms that has been summed on the ^{2}D. The analytes from Table 5.1 have been circled.

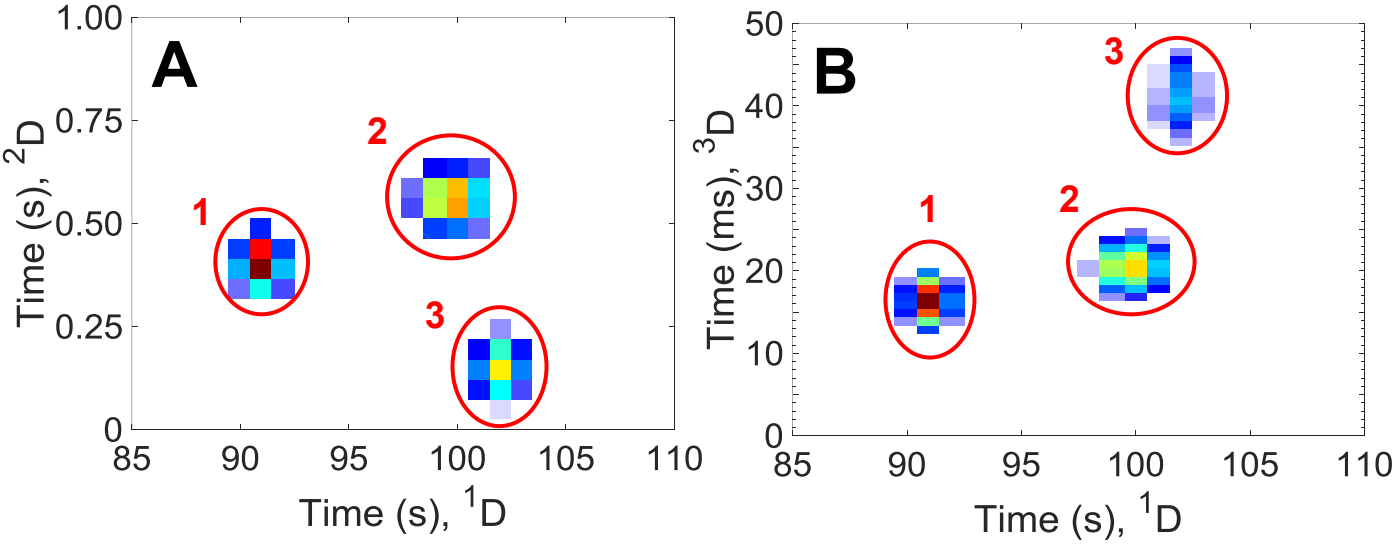


Figure A.5. (A) ^{2}D vs ^{3}D chromatogram of the 115 component test mixture collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{1}D. Because this is the summation of all the ^{1}D modulations, it is difficult to extract meaningful information. Thus (B), (C), and (D) have been prepared to show the benefit of visualizing the data in this manner. (B) Plot of ^{2}D vs ^{3}D chromatogram of the 115 component test mixture summed on ^{1}D between 3 to 4 minutes. (C) Plot of ^{2}D vs ^{3}D chromatogram of the 115 component test mixture summed on ^{1}D between 5 to 6 minutes. (D) Plot of ^{2}D vs ^{3}D chromatogram of the 115 component test mixture summed on ^{1}D between 7 to 8 minutes.

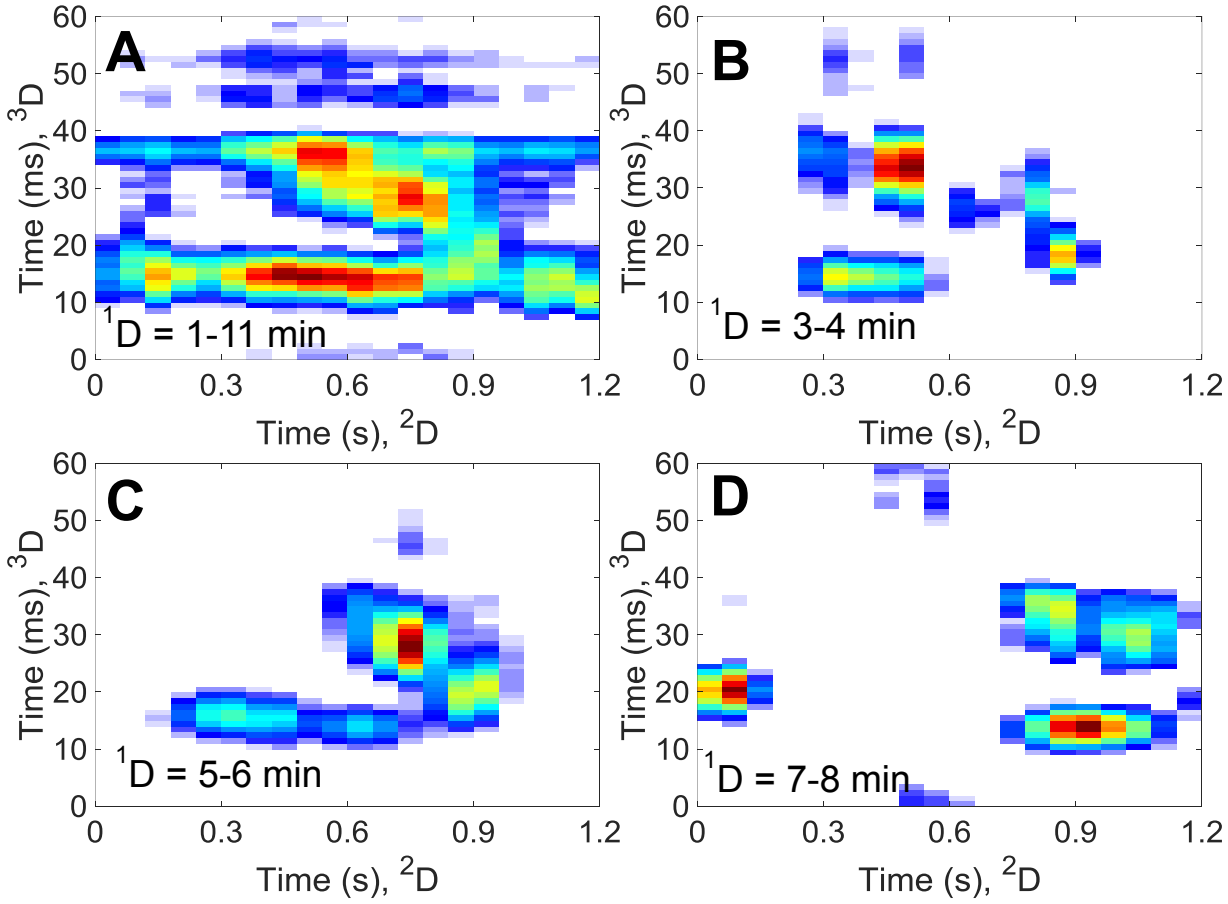


Figure A.6. (A) ^{1}D vs ^{2}D chromatogram of the 115 component test mixture collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{3}D. The analytes from Table 5.2 have been circled. (B) ^{1}D vs ^{3}D chromatogram of the 115 component test mixture collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{2}D. The analytes from Table 5.2 have been circled.

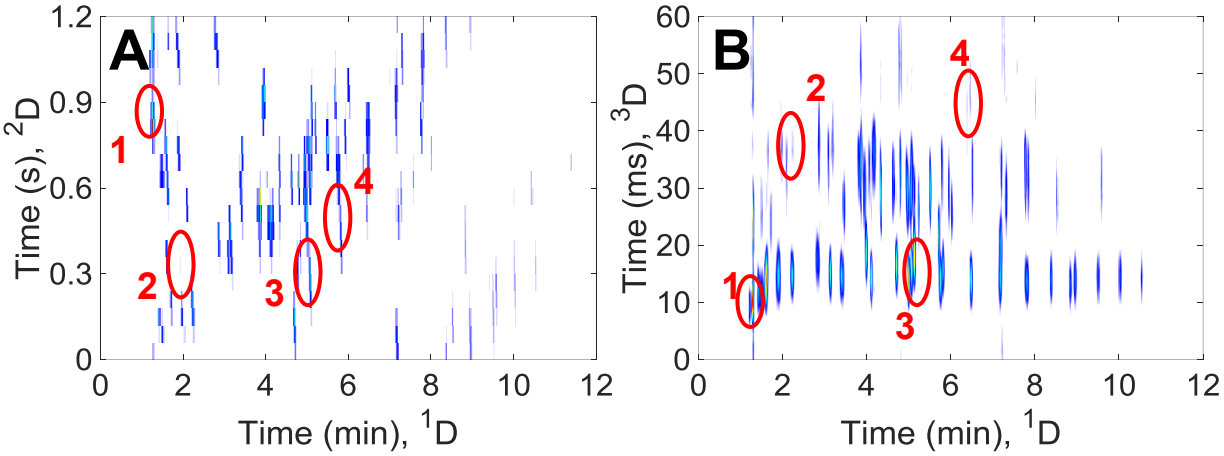
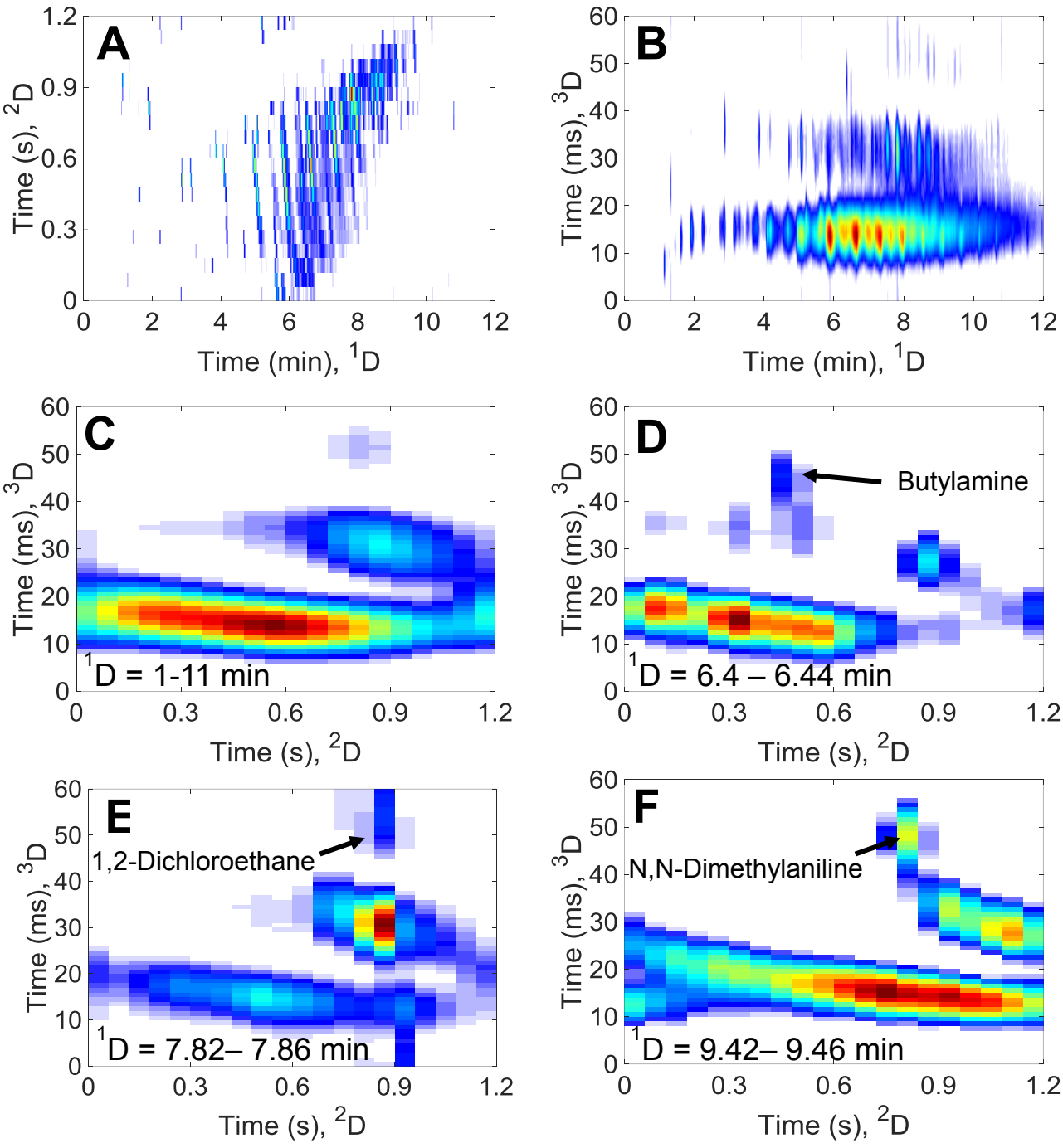


Figure A.7. (A) ^{1}D vs ^{2}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{3}D. (B) ^{1}D vs ^{3}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{2}D. (C) ^{2}D vs ^{3}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{1}D. Diesel fuel is even more complicated than the 115 and 18 component mixtures and because there are thousands of compounds, viewing the data in this manner does not yield much information. Thus (D), (E), and (F) have been prepared where only portions of the ^{1}D chromatogram are shown. In order to improve the visualization, only 3 modulations (3.6 seconds) are shown. (D) ^{2}D vs ^{3}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{1}D from 6.4 to 6.44 minutes. The analyte at a ^{2}D retention time of ~0.45 s and ^{3}D retention time of ~45 ms is butylamine 1 of the 8 spiked analytes. (E) ^{2}D vs ^{3}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{1}D from 7.82 to 7.86 minutes. The analyte at a ^{2}D retention time of ~0.90 s and ^{3}D retention time of ~55 ms is 1,2-dichloroethane 1 of the 8 spiked analytes. (F) ^{2}D vs ^{3}D chromatogram of diesel fuel spiked with 8 compounds collected using a $^{1}P_{M}$ of 1.2 seconds and a $^{2}P_{M}$ of 60 ms that has been summed on the ^{1}D from 9.42 to 9.46 minutes. The analyte at a ^{2}D retention time of ~0.90 s and ^{3}D retention time of ~50 ms is N,N-dimethylaniline 1 of the 8 spiked analytes.


A
Time (s), 2D
Time (min), 1D
B
Time (ms), 3D
Time (min), 1D
C
Time (ms), 3D
1D = 1-11 min
Time (s), 2D
D
Butylamine
Time (ms), 3D
1D = 6.4 – 6.44 min
Time (s), 2D
E
1,2-Dichloroethane
Time (ms), 3D
1D = 7.82– 7.86 min
Time (s), 2D
F
N,N-Dimethylaniline
Time (ms), 3D
1D = 9.42– 9.46 min
Time (s), 2D

Appendix B

Table B.1. List of the 74 kerosene-based fuels analyzed. The fuels span a wide range of distillate cuts and compositional variation. Furthermore, the fuels have different specifications (i.e. types).

Sample Number	Sample Name	Type
1	YA2921HW10	RP-2
2	BG1121GP04	RP-1
3	GRC/0-100 HEP	RP-1
4	WC0721HW01	RP-2
5	LB073009-05	RP-1
6	ZI1521HW10	RP-1
7	CG0721HW10	RP-2
8	LB073009-08	RP-1
9	BB0821HW10	RP-2
10	LB080409-05	RP-1
11	ZI2621HW01	RP-2
12	ZJ1321GP01	RP-2
13	RG3021LS06	UL-RP-1
14	RG3021LS05	RP-TS-5
15	POSF 3327	JP-7
16	POSF 4765	JP-900
17	LB073009-02	RP-1
18	112868	RP-1
19	VI2621LS01	RP-1
20	DB131014	RP-1
21	DC310925	RP-1
22	DC310923	RP-1
23	DB131013	RP-1
24	DB131015	RP-1
25	CL031236	RP-1
26	CB1121HW10	RP-2
27	EA130720	RP-1
28	EB220705	RP-1
29	CHC JP-5	JP-5
30	LB080409-01	RP-1
31	LB073009-01	RP-1
32	A0072256	RP-1
33	CL11-3089	D80
34	LB100413-40	RP-1
35	LB073009-03	RP-1
36	A0072256	RP-1
37	SA1421LS03	UL-RP-1
38	ED060739	RP-1
39	CB1121HW10	RP-2
40	LB073009-09	RP-1
41	LB073009-06	RP-1
42	XC2521HW10	RP-1
43	ZK0821HW20	RP-2
44	ZK2121HW10	RP-2
45	CL11-2928	JP-8
46	CL11-2929	IPK
47	CL11-2654	HRJ
48	CL12-3493	Solvent
49	CA2021HW10	RP-2
50	POSF9641	ATJ-2
51	41910	RP-1
52	B01001634-01	RP-1
53	POSF 4751	JP-8
54	POSF 10359	JP-8
55	POSF 10314	JP-8
56	POSF 10312	JP-8
57	PSOF 10316	JP-8
58	POSF 9326	Jet-A
59	POSF 10358	Jet-A
60	POSF 10311	Jet-A
61	POSF 10315	Jet-A
62	POSF 10369	Jet-A
63	POSF 10313	Jet-A
64	POSF 10337	JP-5
65	POSF 10325	Jet-A
66	POSF 10264	JP-8
67	POSF 10289	JP-5
68	POSF 9698	JP-8
69	EA0721BE01	RP-1
70	CL1621HW10	RP-1
71	DA2321HW20	RP-1
72	DJ2621BE10	RP-1
73	Sample A	RP-1
74	Sample B	RP-1

Table B.2. List of the physical properties for the 74 fuels numbered in Table B.1.

Sample Number	Viscosity 40 °C	Heat of Combustion (btu/lbm)	Hydrogen Content (mass%)	Density 15 °C	Density 45 °C	Density 85 °C
1	1.609	18626	14.202	0.81022	0.78841	0.75906
2	1.696	18626	14.184	0.80879	0.78714	0.75797
3	1.612	18594	14.187	0.81011	0.78801	0.75861
4	1.760	18660	14.368	0.80318	0.78149	0.75241
5	1.710	18555	14.054	0.81362	0.79184	0.76259
6	1.675	18559	14.130	0.81447	0.79259	0.76330
7	1.645	18583	14.215	0.80999	0.78814	0.75887
8	1.600	18580	14.183	0.80498	0.78308	0.75378
9	1.705	18556	14.178	0.80890	0.78711	0.75786
10	1.663	18587	14.160	0.81318	0.79139	0.76210
11	1.852	18603	14.276	0.80822	0.78667	0.75770
12	1.662	18608	14.218	0.80864	0.78690	0.75768
13	1.593	18604	14.242	0.80743	0.78562	0.75630
14	1.584	18602	14.216	0.80745	0.78559	0.75622
15	1.537	18690	14.536	0.79124	0.76944	0.74008
16	1.889	18289	13.073	0.86994	0.84793	0.81854
17	1.588	18568	14.055	0.81131	0.78939	0.76003
18	1.472	18590	14.125	0.80520	0.78324	0.75375
19	1.593	18598	14.210	0.80962	0.78769	0.75832
20	2.324	18683	14.451	0.80694	0.78587	0.75771
21	1.600	18619	14.234	0.80672	0.78486	0.75549
22	1.623	18623	14.211	0.80762	0.78579	0.75647
23	1.647	18644	14.249	0.80738	0.78554	0.75634
24	2.117	18489	13.790	0.83545	0.81387	0.78472
25	1.031	18607	14.215	0.80329	0.78066	0.74973
26	1.688	18546	14.122	0.81415	0.79237	0.76320
27	1.652	18514	13.966	0.81743	0.79545	0.76616
28	1.559	18538	14.005	0.81227	0.79032	0.76088
29	1.347	18594	14.117	0.80160	0.77923	0.74923
30	1.654	18553	14.179	0.81103	0.78927	0.76001
31	1.720	18559	14.146	0.81583	0.79401	0.76494
32	1.433	18636	14.057	0.80272	0.78069	0.75090
33	1.635	18668	14.570	0.79370	0.77212	0.74289
34	1.644	18582	14.093	0.80887	0.78719	0.75782
35	1.652	18590	14.157	0.81352	0.79178	0.76240
36	1.593	18563	14.188	0.81222	0.79030	0.76096

37	1.603	18584	14.213	0.80990	0.78809	0.75878
38	1.667	18497	13.970	0.81794	0.79616	0.76695
39	1.694	18526	14.157	0.81424	0.79244	0.76325
40	1.608	18552	14.160	0.81269	0.79075	0.76135
41	1.726	18592	14.266	0.80938	0.78767	0.75852
42	1.588	18582	14.253	0.80924	0.78736	0.75796
43	1.661	18550	14.184	0.81349	0.79164	0.76243
44	1.526	18550	14.180	0.81250	0.79057	0.76083
45	1.344	18487	13.779	0.80371	0.78144	0.75137
46	1.128	18815	15.192	0.75995	0.73768	0.70713
47	1.522	18819	15.097	0.76476	0.74310	0.71342
48	1.989	18188	12.944	0.88432	0.86154	0.83123
49	1.645	18572	14.188	0.81090	0.78919	0.75989
50	1.481	18797	15.247	0.75843	0.73674	0.70738
51	1.512	18602	14.266	0.80469	0.78283	0.75327
52	1.620	18579	14.263	0.80732	0.78554	0.75633
53	1.340	18394	13.932	0.80370	0.78150	0.75150
54	1.330	18553	13.988	0.80240	0.78030	0.75020
55	1.510	18462	13.908	0.81240	0.79050	0.76080
56	1.370	18463	13.751	0.81140	0.78930	0.75940
57	1.290	18522	13.971	0.80480	0.78250	0.75220
58	1.190	18483	13.761	0.80500	0.78240	0.75190
59	1.430	18459	13.946	0.80550	0.78440	0.75470
60	1.370	18450	13.499	0.82010	0.79790	0.76790
61	1.140	18623	14.282	0.79180	0.76970	0.73960
62	1.180	18533	13.985	0.79730	0.77480	0.74430
63	1.480	18488	13.618	0.81020	0.78830	0.75900
64	1.370	18514	13.852	0.80800	0.78580	0.75590
65	1.310	18506	13.920	0.80320	0.78100	0.75090
66	1.140	18584	14.410	0.77990	0.75740	0.72860
67	1.570	18429	13.580	0.82680	0.80480	0.77540
68	1.300	18530	14.078	0.79920	0.77690	0.74670
69	1.680	18611	14.246	0.80962	0.78735	0.75821
70	1.601	18610	14.289	0.80667	0.78404	0.75473
71	1.602	18627	14.260	0.80684	0.78492	0.75556
72	2.148	18508	13.644	0.83641	0.81429	0.78558
73	1.563	18624	14.441	0.80416	0.78234	0.75298
74	1.586	18625	14.436	0.80608	0.78435	0.75520

Figure B.1. (A) Representative GC × GC – TOFMS TIC chromatogram for the fuel sample set (Sample 69: POSF 10289) with the three major chemical classes highlighted: alkanes (paraffins), cycloalkanes (cycloparaffins), and aromatics. (B) One dimensional GC separation of Sample 69. This figure was prepared by summing the second dimension separations of the GC × GC TIC in (A) onto the first dimension axis.

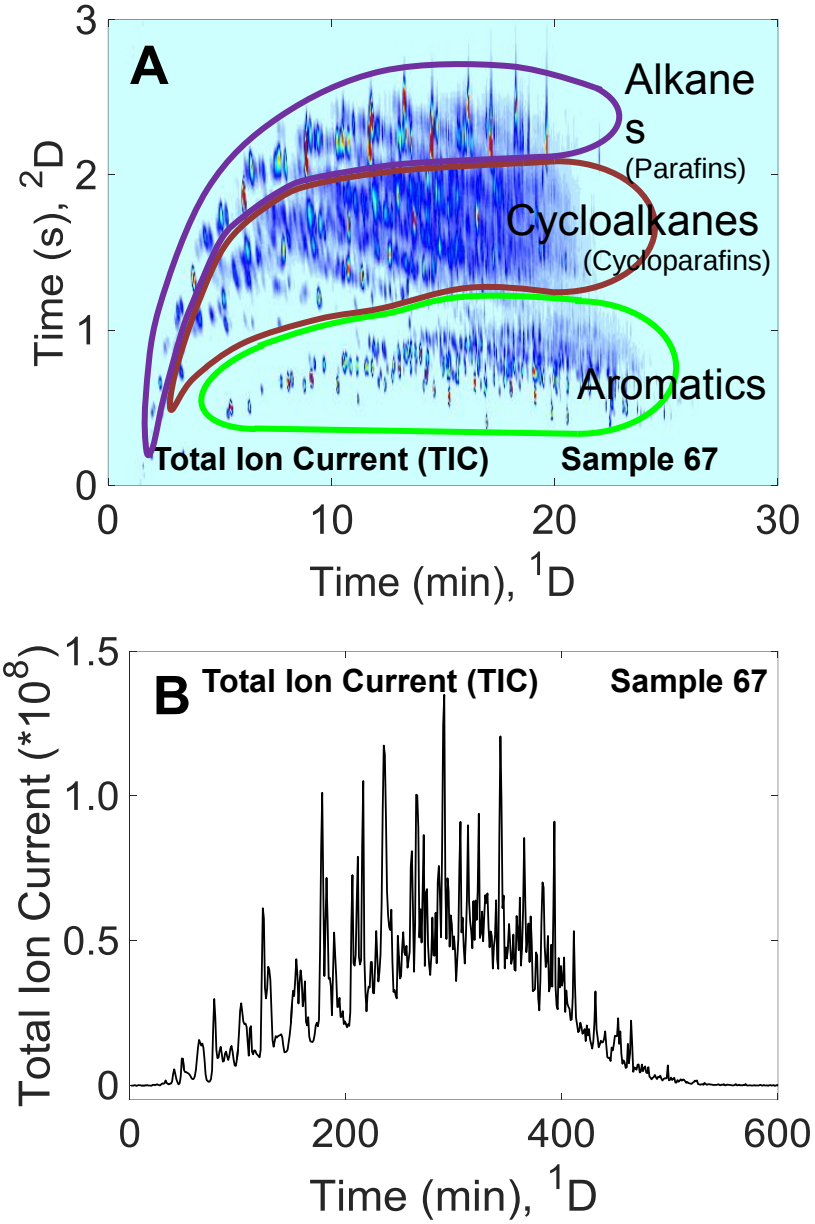


Figure B.2. Analytical ion chromatograms (AIC) for Sample 67 showing the ability to use selective *m/z* to identify the chemical compound classes (i.e., monocycloalkanes, dicycloalkanes, alkyl benzenes, etc.). Within each figure is a mass spectrum of an arbitrarily selected analyte compound which has been circled. (A) AIC of Sample 69 for alkanes. (B) AIC of Sample 69 for monocycloalkanes. (C) AIC of Sample 69 for dicycloalkanes. (D) AIC of Sample 69 for alkyl benzenes (E) AIC of Sample 69 for naphthalenes. (F) AIC of Sample 69 for tetralins and indans.

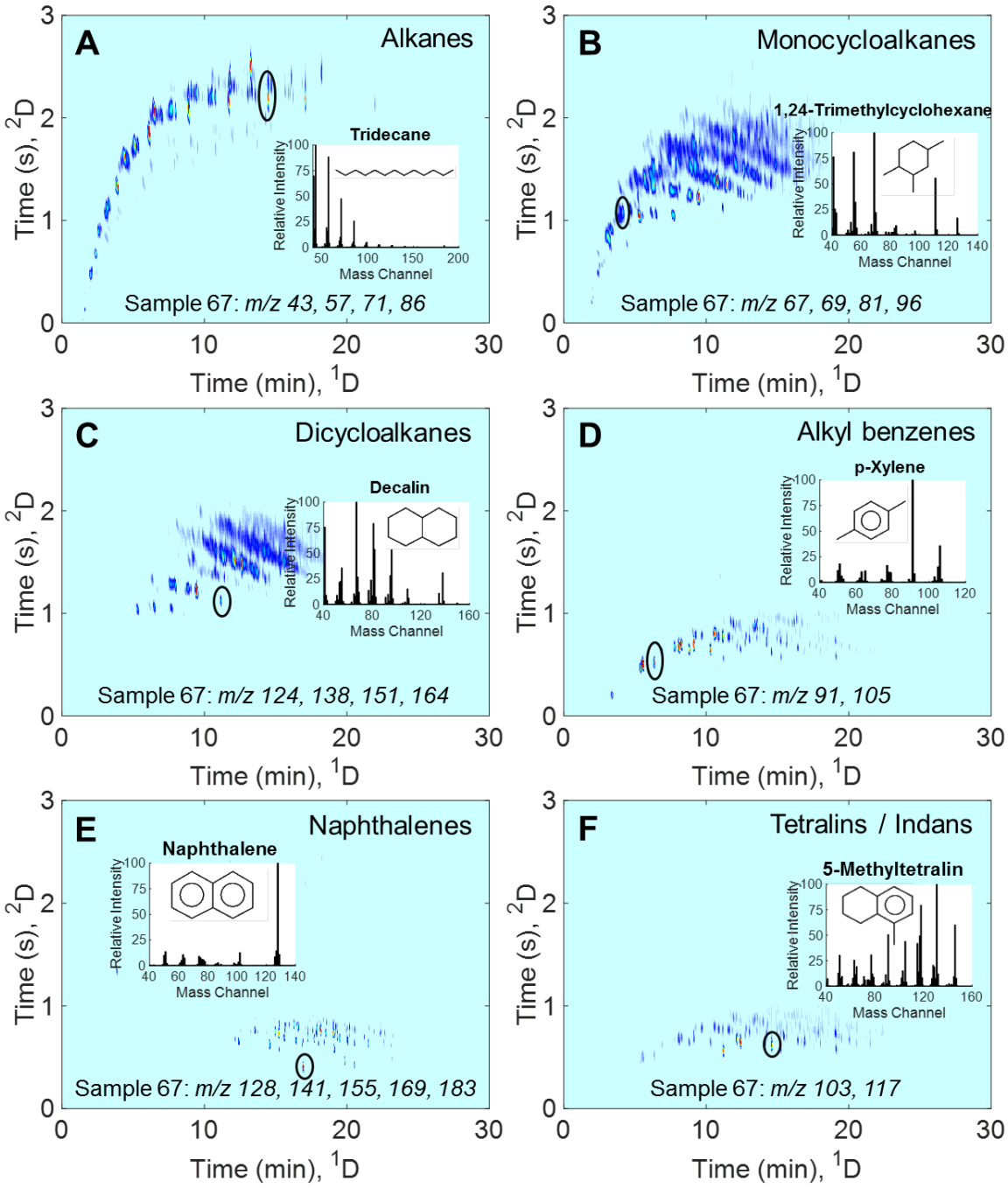
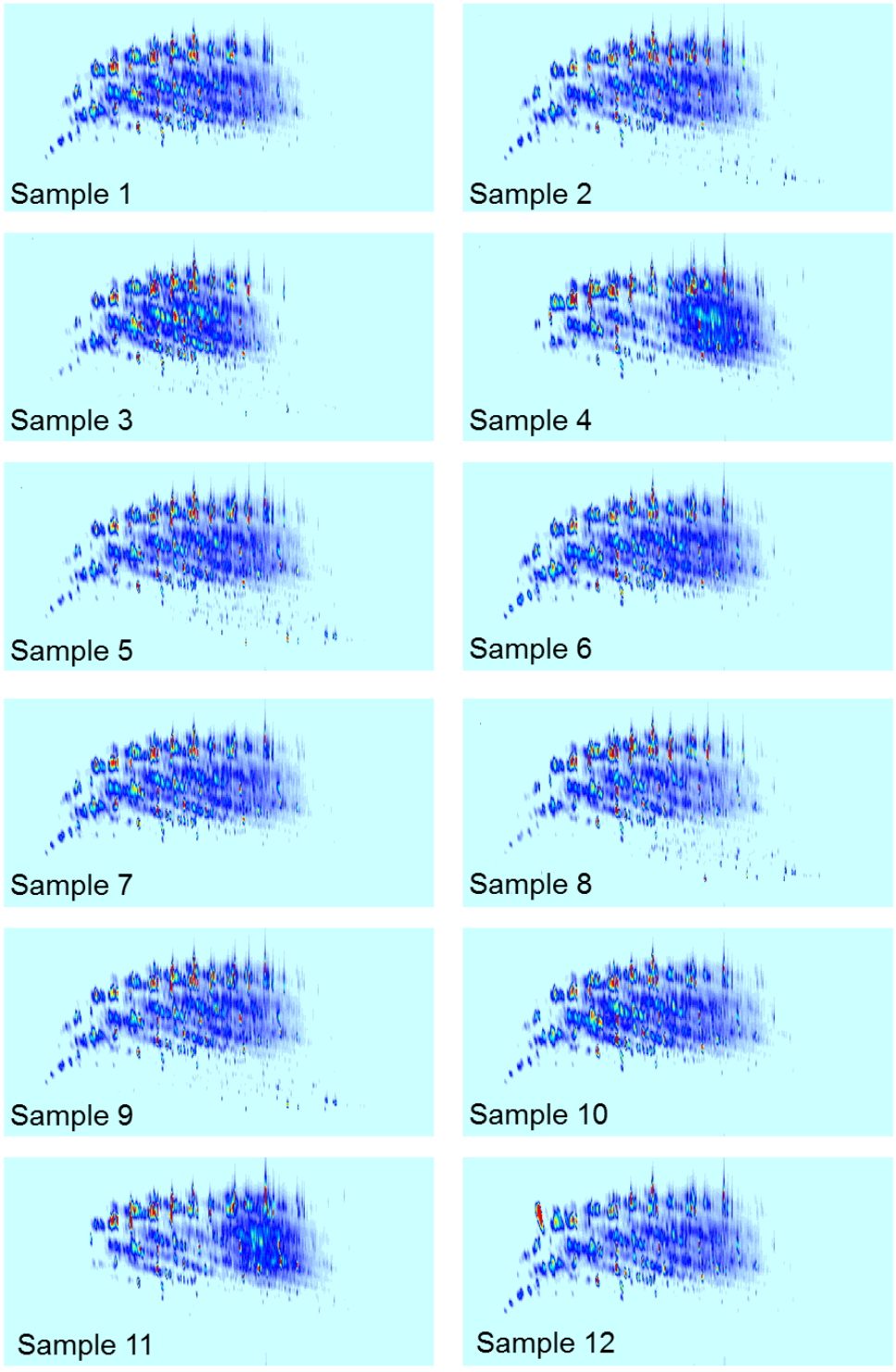
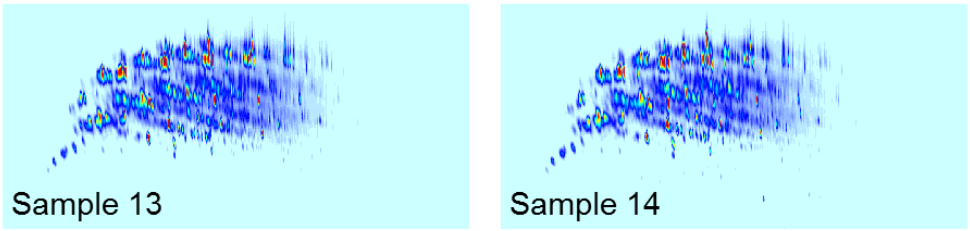
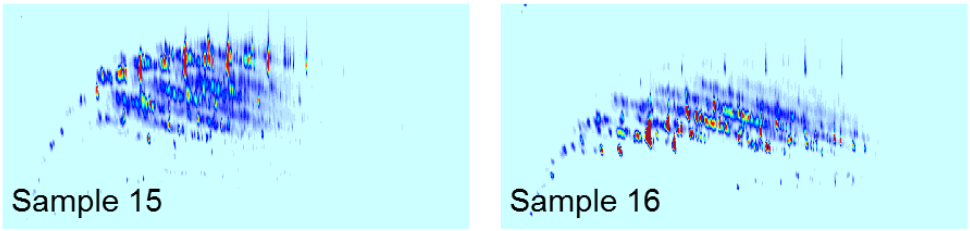
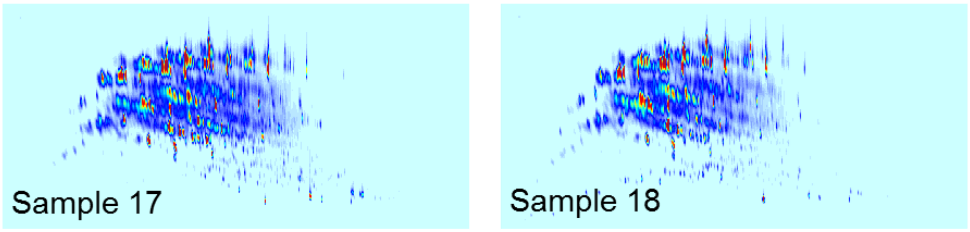


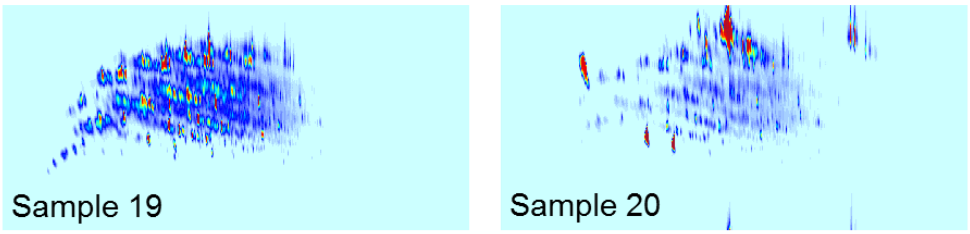
Figure B.3. GC × GC – TOFMS chromatograms of all 74 fuel samples. There are 9 samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) that are dissimilar in their chemical fingerprint from the majority of the samples.

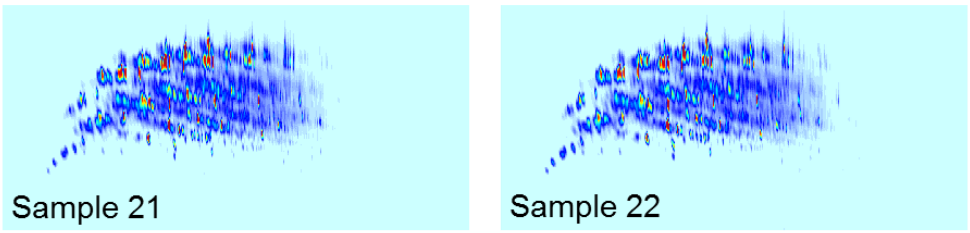



Sample 13
Sample 14

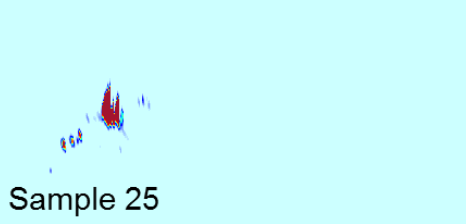

Sample 15
Sample 16

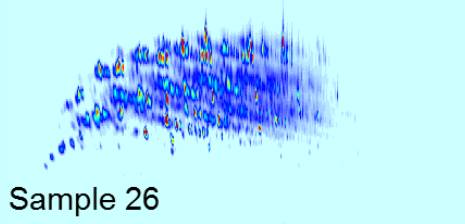

Sample 17
Sample 18

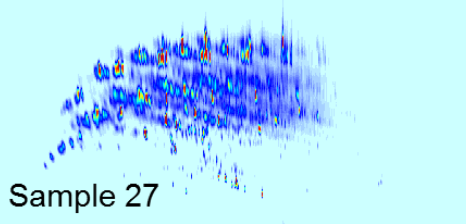

Sample 19
Sample 20

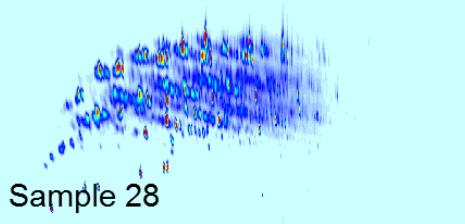

Sample 21
Sample 22

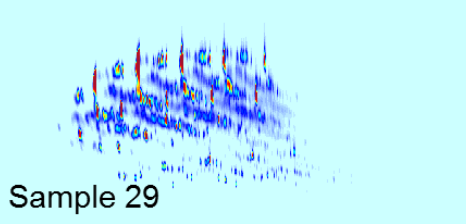

Sample 23
Sample 24

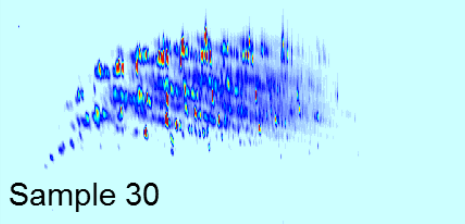

Sample 25

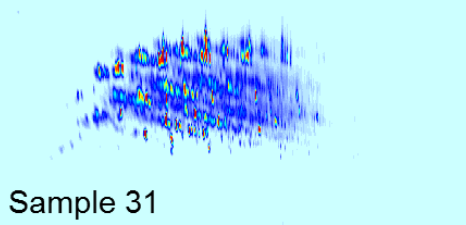

Sample 26

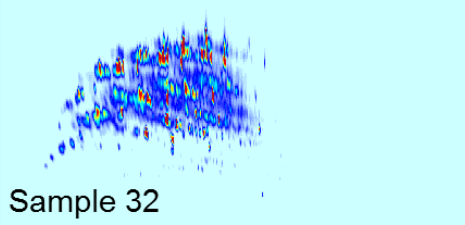

Sample 27

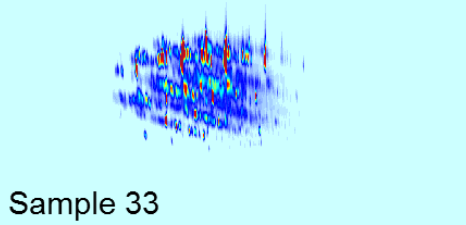

Sample 28

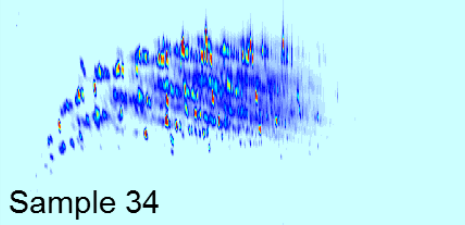

Sample 29

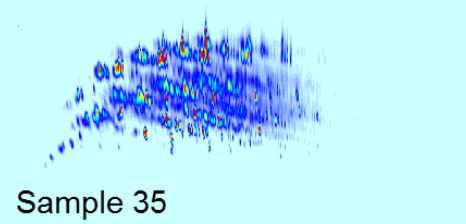

Sample 30

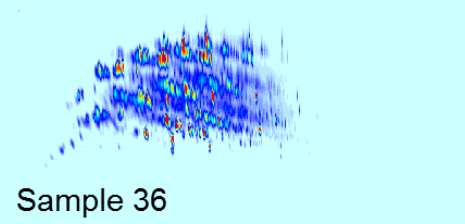

Sample 31

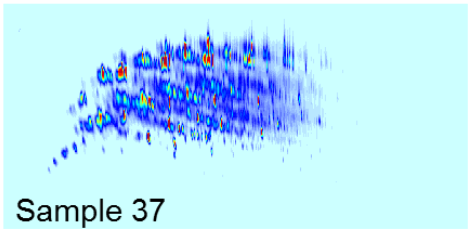

Sample 32

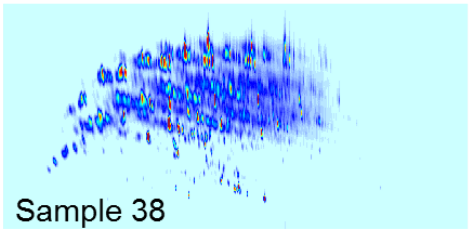

Sample 33

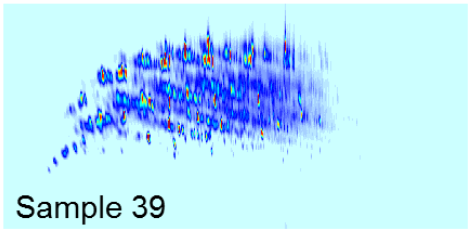

Sample 34

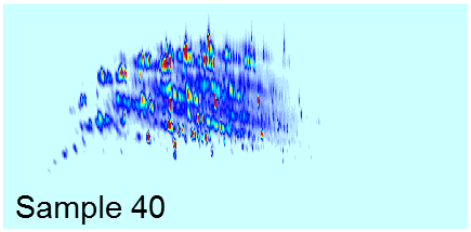

Sample 35

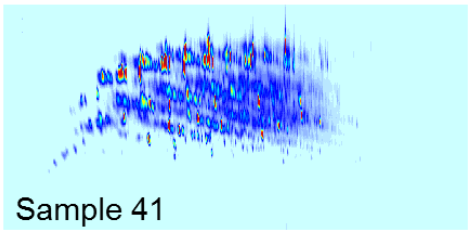

Sample 36

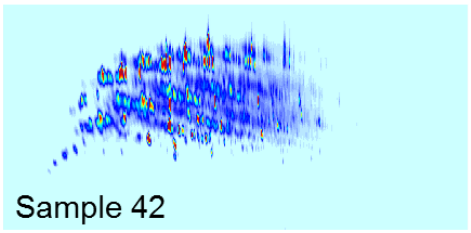

Sample 37

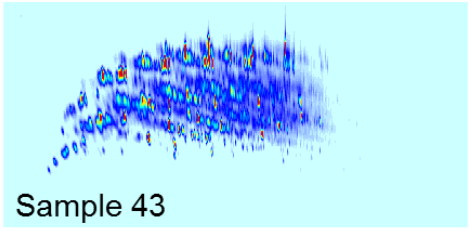

Sample 38

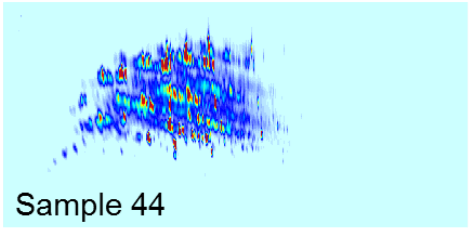

Sample 39

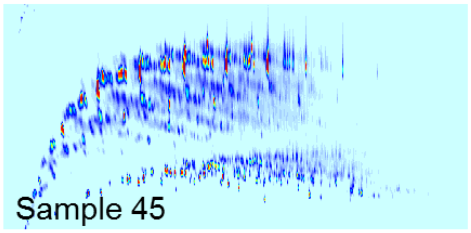

Sample 40

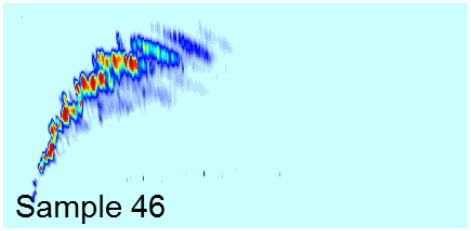

Sample 41

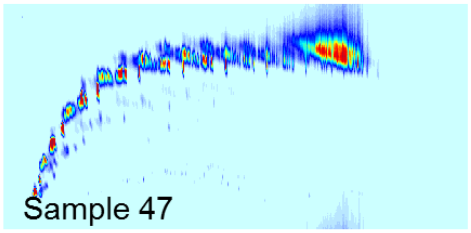

Sample 42

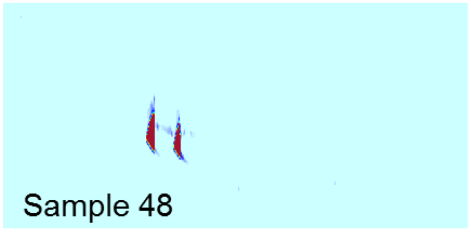

Sample 43

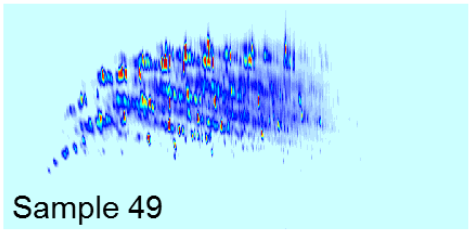

Sample 44

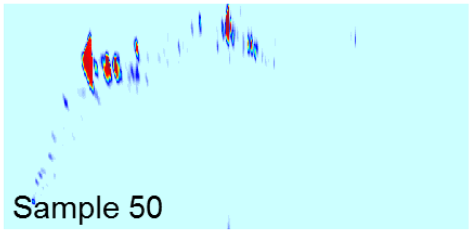

Sample 45

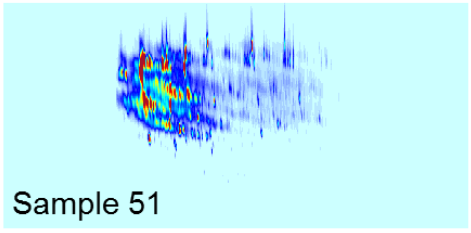

Sample 46

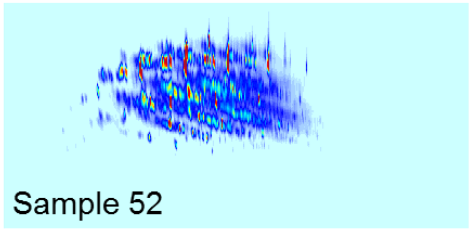

Sample 47

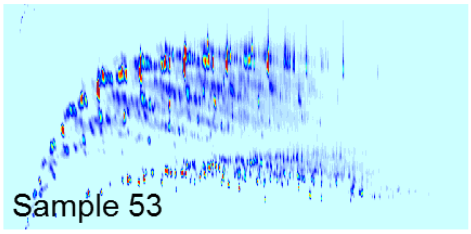

Sample 48

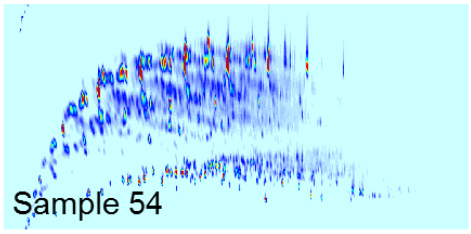

Sample 49

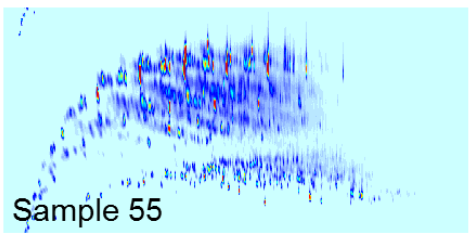

Sample 50

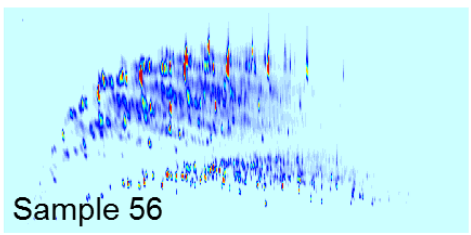

Sample 51

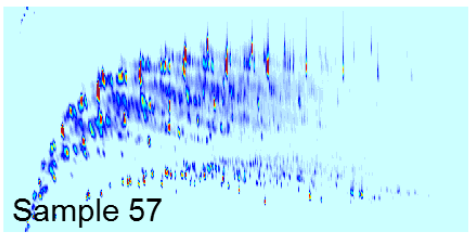

Sample 52

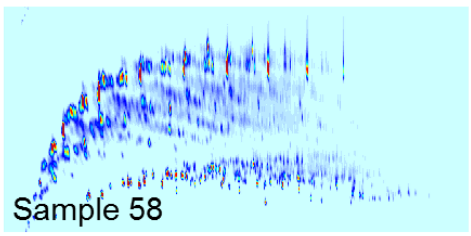

Sample 53

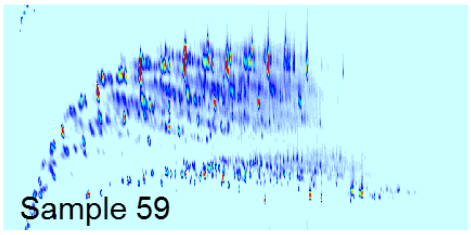

Sample 54

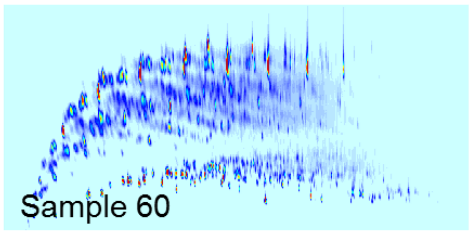

Sample 55

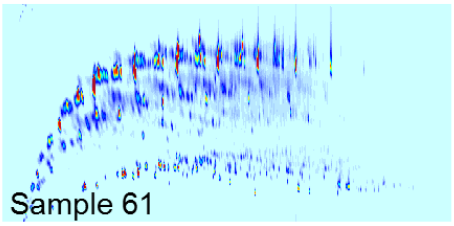

Sample 56

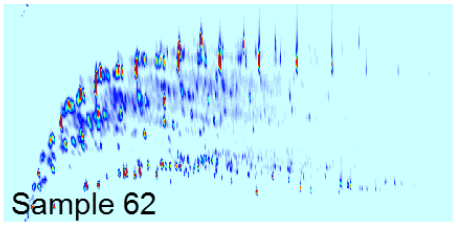

Sample 57

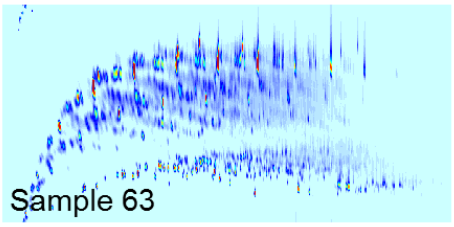

Sample 58

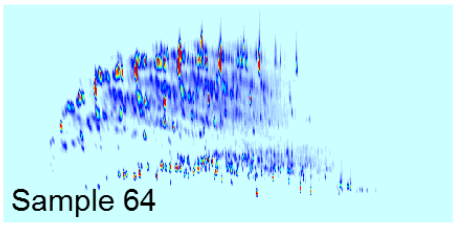

Sample 59

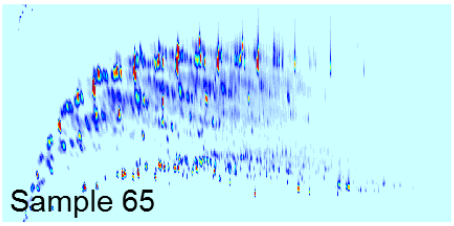

Sample 60

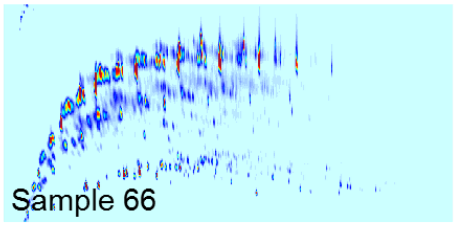

Sample 61

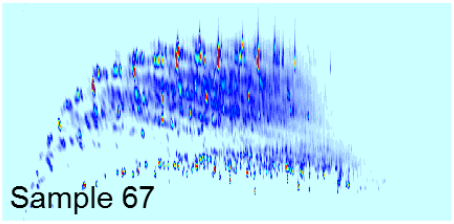

Sample 62

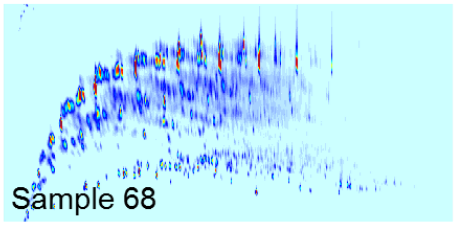

Sample 63

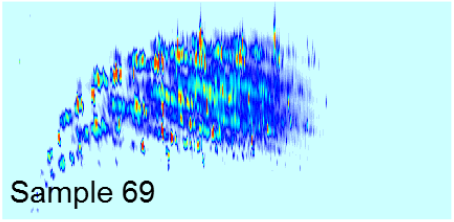

Sample 64

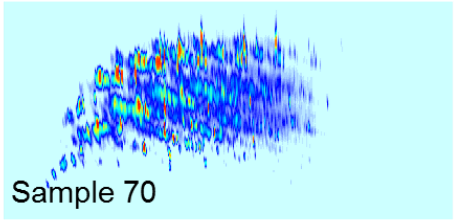

Sample 65

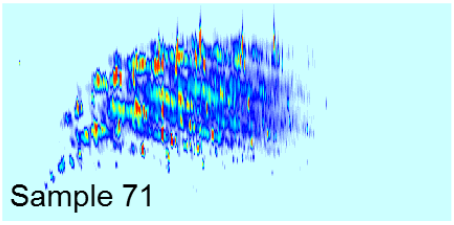

Sample 66

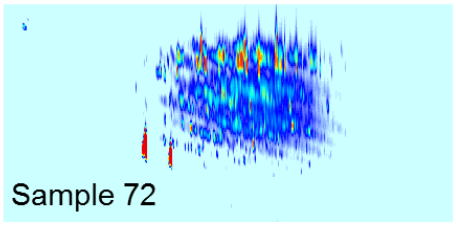

Sample 67


Sample 68


Sample 69


Sample 70


Sample 71


Sample 72

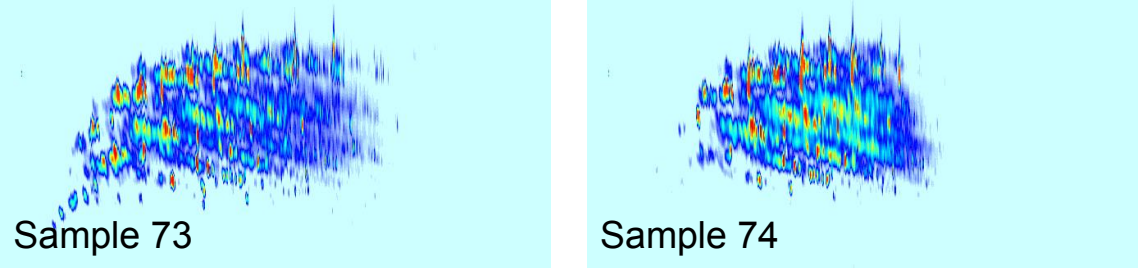


Figure B.4. Q residuals reduced vs the hoteling T^2 reduced for the PLS prediction of viscosity using all of the GC × GC – TOFMS chromatograms (see Figure 7.2). The PLS model used six latent variables (LVs) and had an NRMSECV error of ~ 11% for both replicates. While the calibration plot in Figure 7.2 had several outliers (Samples 16, 20, 24, 25, 48, 50, and 72), the rest of the samples were reasonably close to the ideal calibration line (slope of one) indicating an overall good model. Evaluation of the linear regression vectors (LRVs), indicated that there were several compounds that were dominating the model which should not be the case since viscosity is a bulk property and should be based on a large number of compounds. Thus, to evaluate the model, a Q residuals and hotelling T^2 reduced plot was prepared. The hotelling T^2 reduced values describe how different samples are away from the model. In other words, how chemically different are the samples from the typical sample. The Q residual reduced values describe the goodness-of-fit or how well each sample is described by the model. We chose to exclude samples that had a high hoteling T^2 value (Samples 16, 20, 24, 25, 46, 47, 48, 50, and 72). The purples lines is the Hotelling T^2 values that was chosen as the threshold for a fuel to be included. There were a few samples that had modestly elevated Q residuals reduced, but low hoteling T^2 reduced values. It was elected to keep those samples in the model. This resulted in a ~1% increase in NRMSECV of the models. The 9 samples that were excluded are significantly different than the typical sample. For example, Sample 48 (see Figure B.3) is only composed of 2 compounds that are grossly overloaded with some additional minor compounds. Modeling of this sample is not possible.

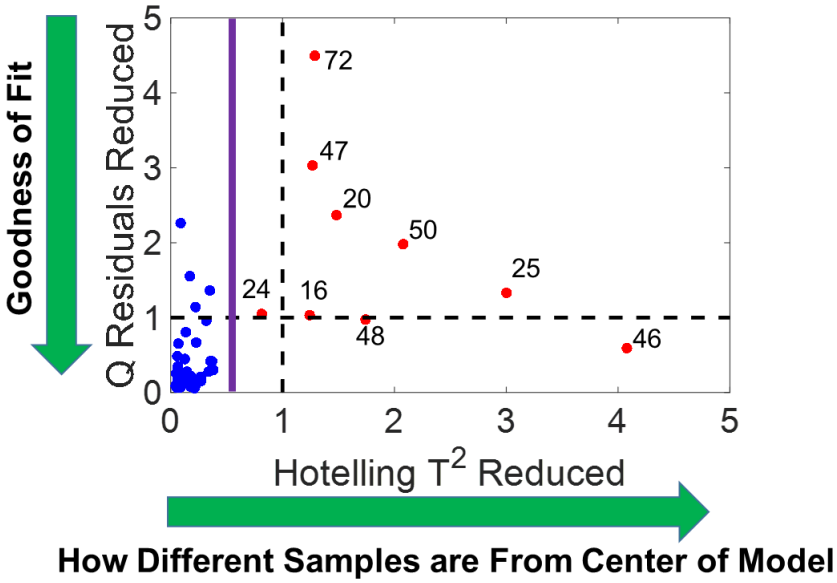


Figure B.5. PLS Prediction of viscosity using the GC × GC – TOFMS chromatograms excluding Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 which were found to have large hoteling T^2 values. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 6% for both replicates

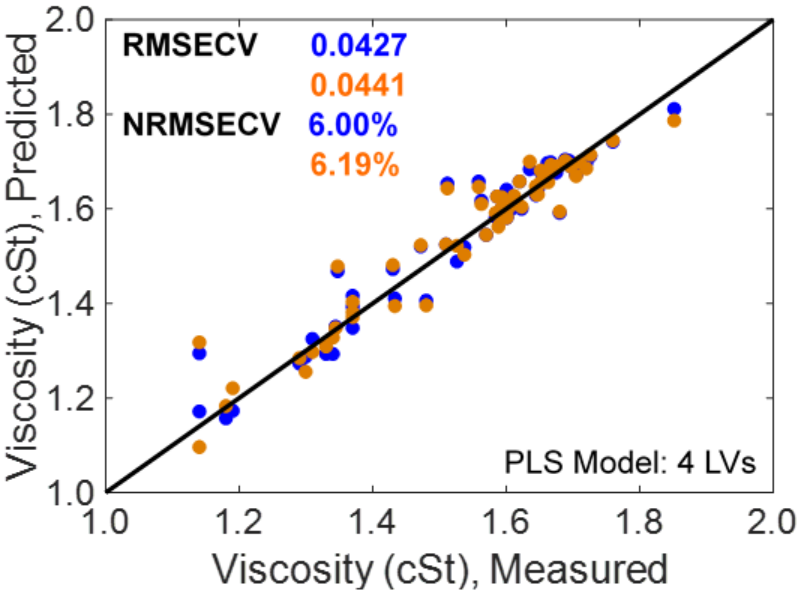


Figure B.6. Heat of combustion corrected for hydrogen content.

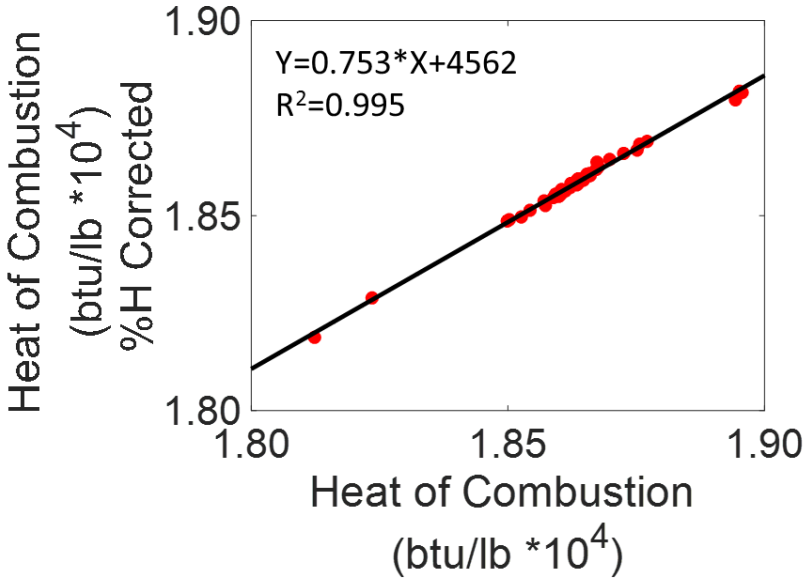


Figure B.7. (A) PLS Prediction of heat of combustion using all the GC × GC – TOFMS chromatograms. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 11% for both replicates. Similar to the viscosity, there are several samples that are not close to the ideal line. Evaluation of the Q residuals reduced and hoteling T^2 reduced (not shown for brevity) indicated that the same samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) exhibited high hoteling T^2 values. (B) Linear regression vectors (LRVs) for the PLS evaluation of heat of combustion using all the samples. There are several compounds that dominate the model. This is consistent with LRVs seen in the viscosity modeling when all samples were included. (C) PLS modeling of the heat of combustion with (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 excluded. The error is ~13% for both replicates. It should be noted that, the method precision (repeatability) and bias for ASTM D4809 are 0.096 and 0.089 MJ/kg (~ 40 btu/lb) respectively, and the model quality based on true values may therefore be better than indicated.

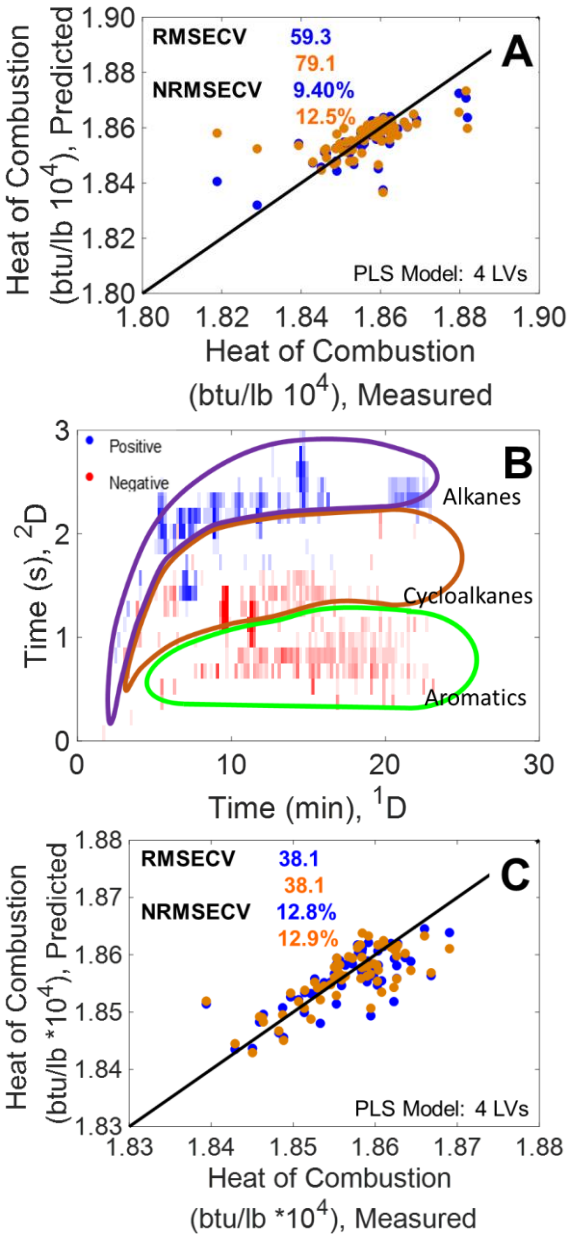


Figure B.8. (A) PLS Prediction of hydrogen content using all the GC × GC – TOFMS chromatograms. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 11% for both replicates. Similar to the viscosity and heat of combustion, there are several samples that are not close to the ideal line. While the PLS modeling of the hydrogen content yielded a better calibration curve, the evaluation of the Q residuals reduced and hoteling T^2 reduced (not shown for brevity) indicated that the same samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) exhibited high hoteling T^2 values. (B) Linear regression vectors (LRVs) for the PLS evaluation of hydrogen content using all the samples. There are several compounds that dominate the model. This is consistent with LRVs seen in the viscosity and heat of combustion modeling when all samples were included. (C) PLS modeling of the hydrogen content with (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 excluded. The error is ~10% for both replicates.

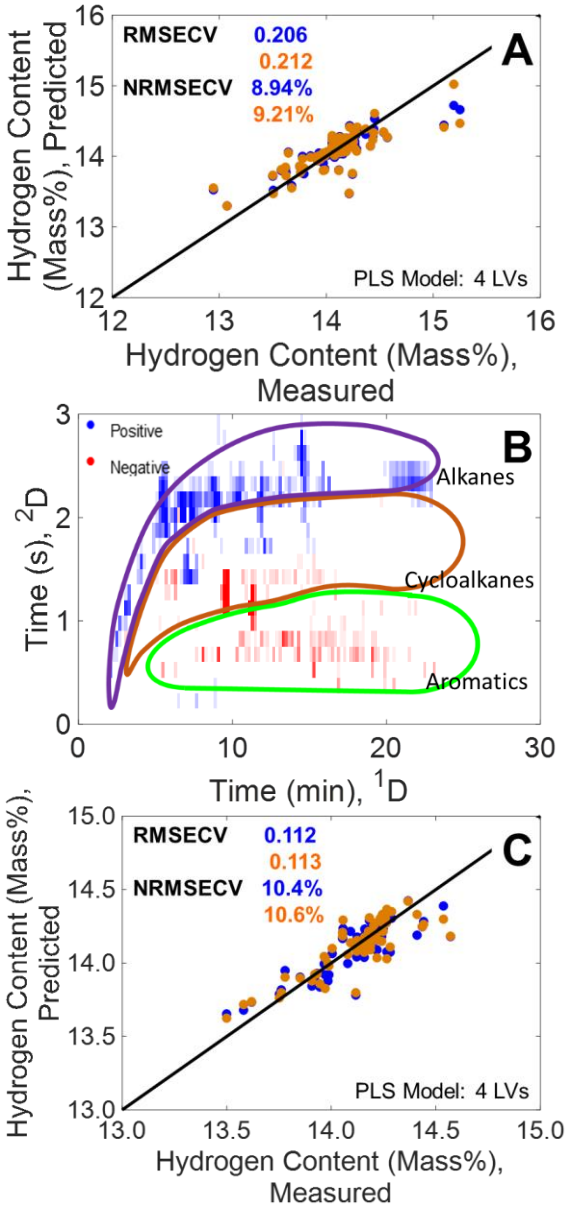


Figure B.9. (A) PLS Prediction of density at 15 °C using all the GC × GC – TOFMS chromatograms. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 9% for both replicates. Similar to the viscosity, heat of combustion, and hydrogen content, there are several samples that are not close to the ideal line. The evaluation of the Q residuals reduced and hoteling T^2 reduced (not shown for brevity) indicated that the same samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) exhibited high hoteling T^2 values. (B) Linear regression vectors (LRVs) for the PLS evaluation of density at 15 °C using all the samples. There are several compounds that dominate the model. This is consistent with LRVs seen in the viscosity, heat of combustion, and hydrogen content modeling when all samples were included. (C) PLS modeling of the density at 15 °C with (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 excluded. The error is ~8% for both replicates.

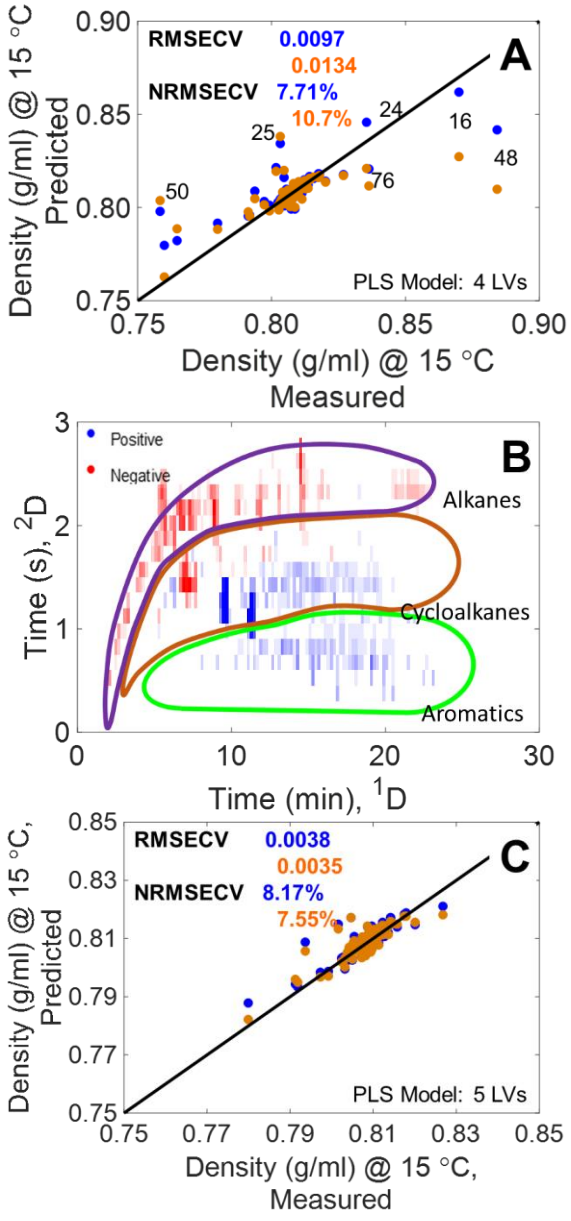


Figure B.10. (A) PLS Prediction of density at 45 °C using all the GC × GC – TOFMS chromatograms. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 9% for both replicates. Similar to the viscosity, heat of combustion, and hydrogen content, there are several samples that are not close to the ideal line. The evaluation of the Q residuals reduced and hoteling T^2 reduced (not shown for brevity) indicated that the same samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) exhibited high hoteling T^2 values. (B) Linear regression vectors (LRVs) for the PLS evaluation of density at 45 °C using all the samples. There are several compounds that dominate the model. This is consistent with LRVs seen in the viscosity, heat of combustion, and hydrogen content modeling when all samples were included. (C) PLS modeling of the density at 45 °C with (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 excluded. The error is ~8% for both replicates.

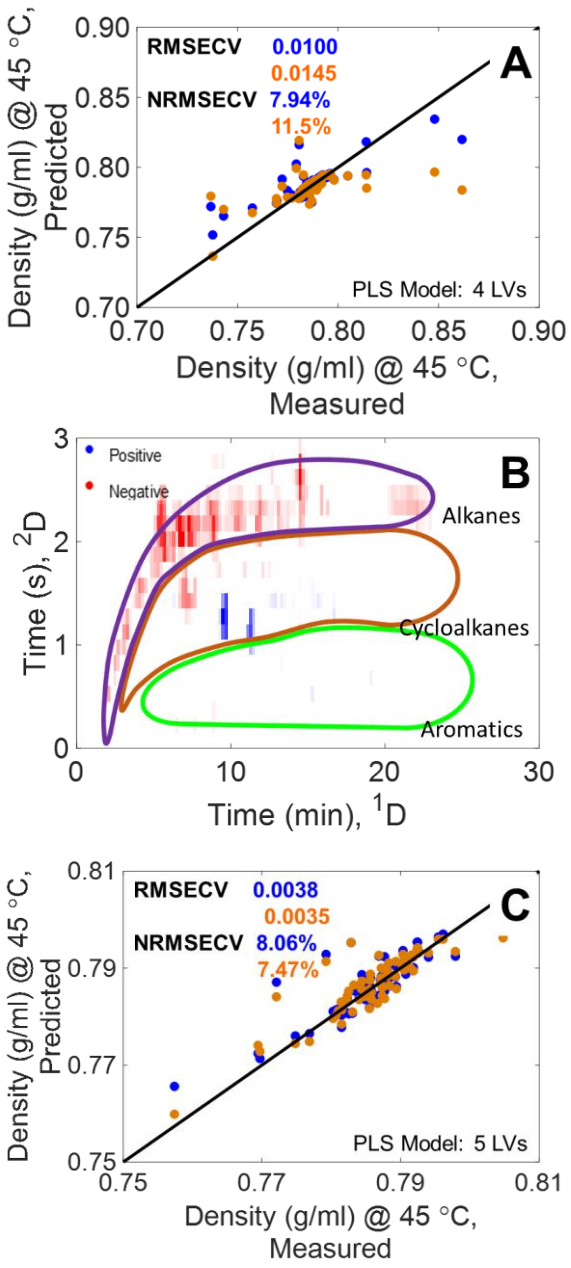
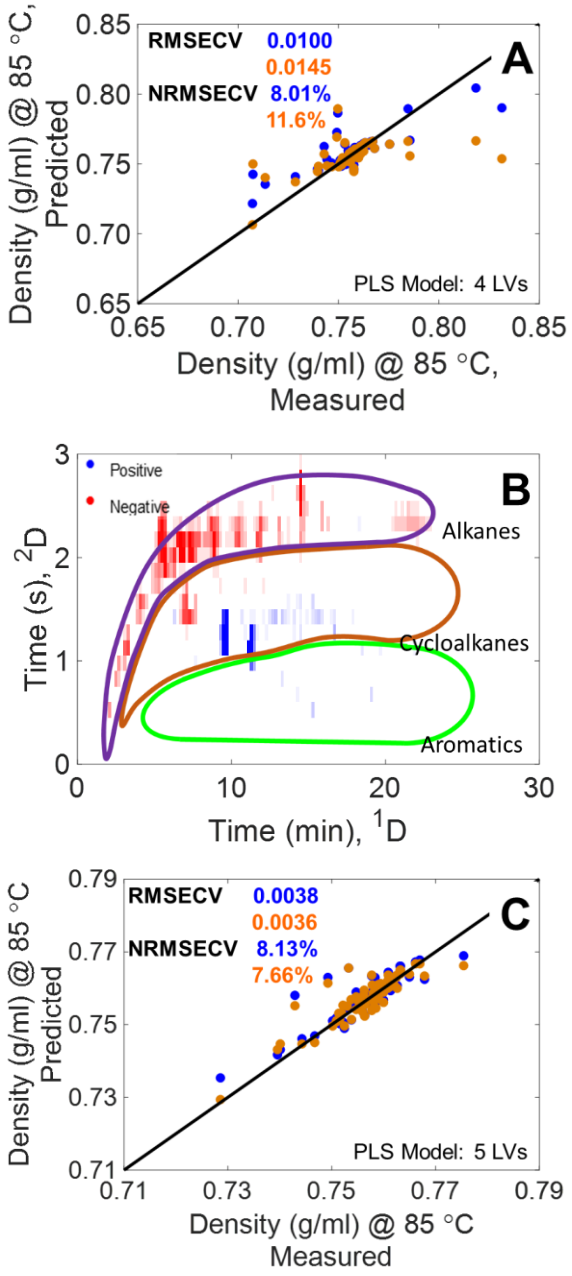


Figure B.11. (A) PLS Prediction of density at 85 °C using all the GC × GC – TOFMS chromatograms. The PLS model used four latent variables (LVs) and had an NRMSECV error of ~ 9% for both replicates. Similar to the viscosity, heat of combustion, and hydrogen content, there are several samples that are not close to the ideal line. The evaluation of the Q residuals reduced and hoteling T^2 reduced (not shown for brevity) indicated that the same samples (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72) exhibited high hoteling T^2 values. (B) Linear regression vectors (LRVs) for the PLS evaluation of density at 85 °C using all the samples. There are several compounds that dominate the model. This is consistent with LRVs seen in the viscosity, heat of combustion, and hydrogen content modeling when all samples were included. (C) PLS modeling of the density at 85 °C with (Samples 16, 20, 24, 25, 46, 47, 48, 60, and 72 excluded. The error is ~8% for both replicates.



Appendix C

Figure C.1. Compact Rapid Assessment of Fuel Thermal Integrity (CRAFTI) instrument. (A) Overview of entire instrument. (B) Enhanced view of the test article. (C) View of the thermocouples placed on the test article.

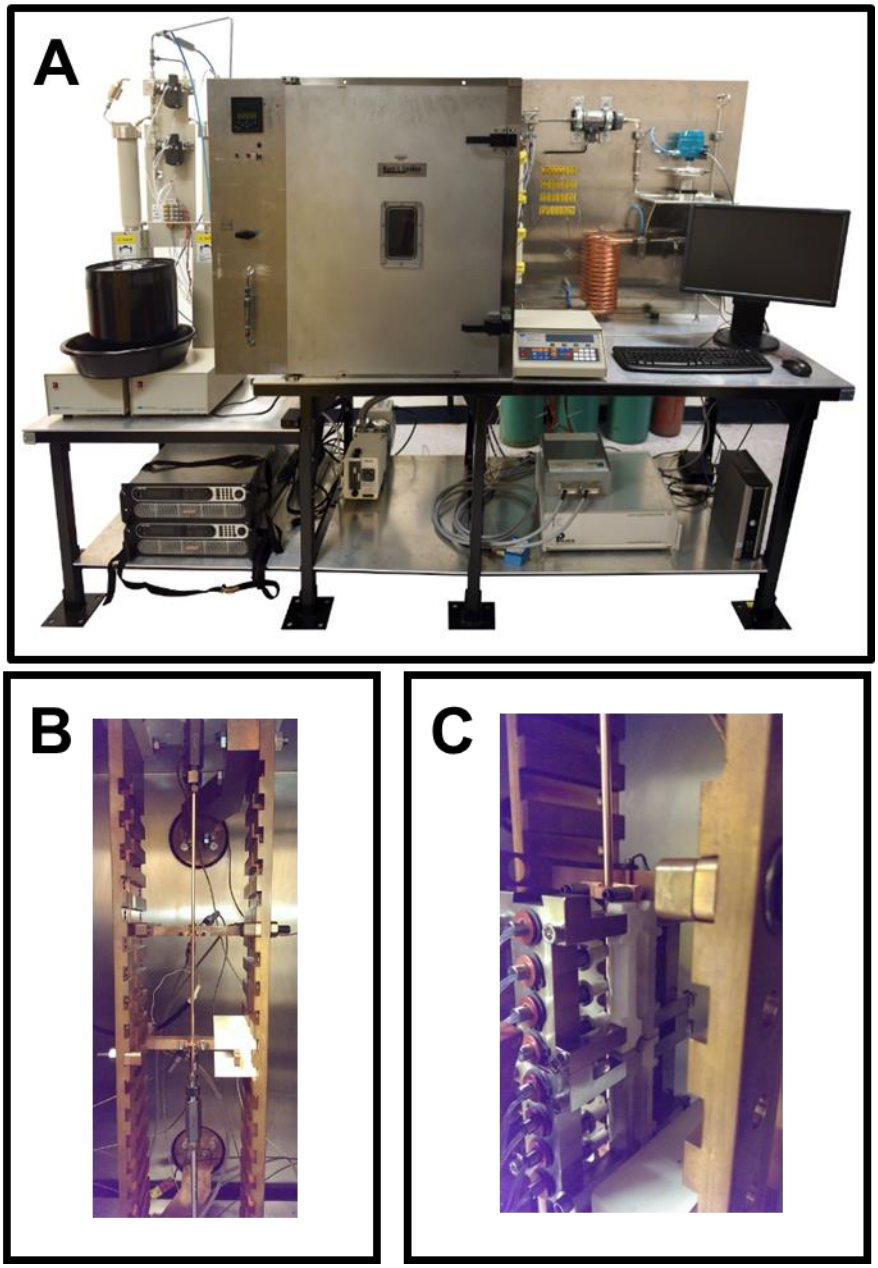


Figure C.2. Two forms of carbon that best correlated with ΔP. (A) ΔP as a function of chemisorbed carbon in the heated zone (B) ΔP as a function of logarithmically scaled chemisorbed carbon in the heated zone. (C) ΔP as a function of amorphous carbon in the exit zone.

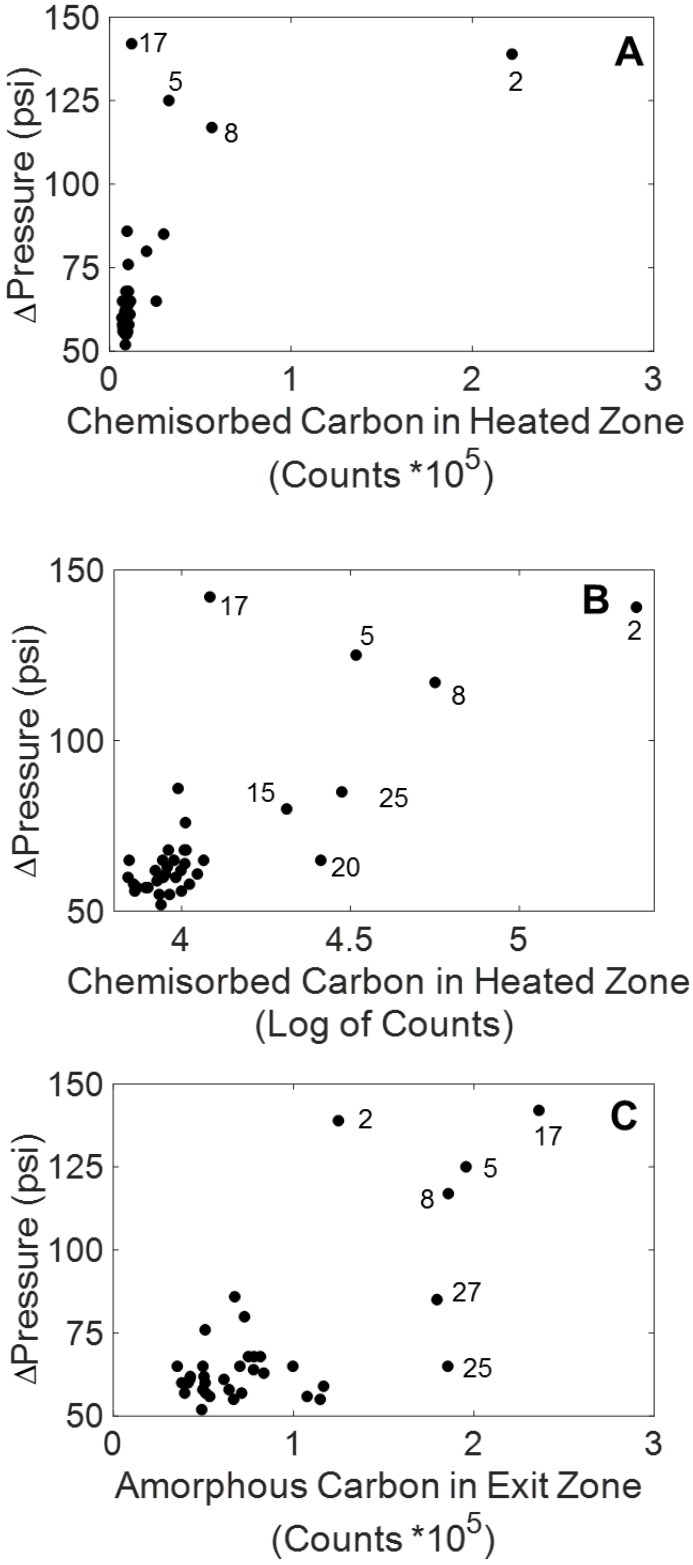


Figure C.3. (A) PLS prediction of ΔPressure using the entire chromatogram for the 36 samples. The model yields an error of ~17% but more troublesome is that some fuels have extremely high predicted ΔP when their measured ΔP is low. (B) The linear regression vectors for the PLS prediction of ΔP. The aromatics are the major species that are attributed to increasing the ΔP. The n-alkanes seem to also correlate with increasing ΔP. The cycloalkanes and the iso-alkanes are negatively correlated with increasing ΔP. The PLS model was deemed insufficient to predict the ΔP so a feature selection technique was implemented prior to the PLS modeling of the ΔP.

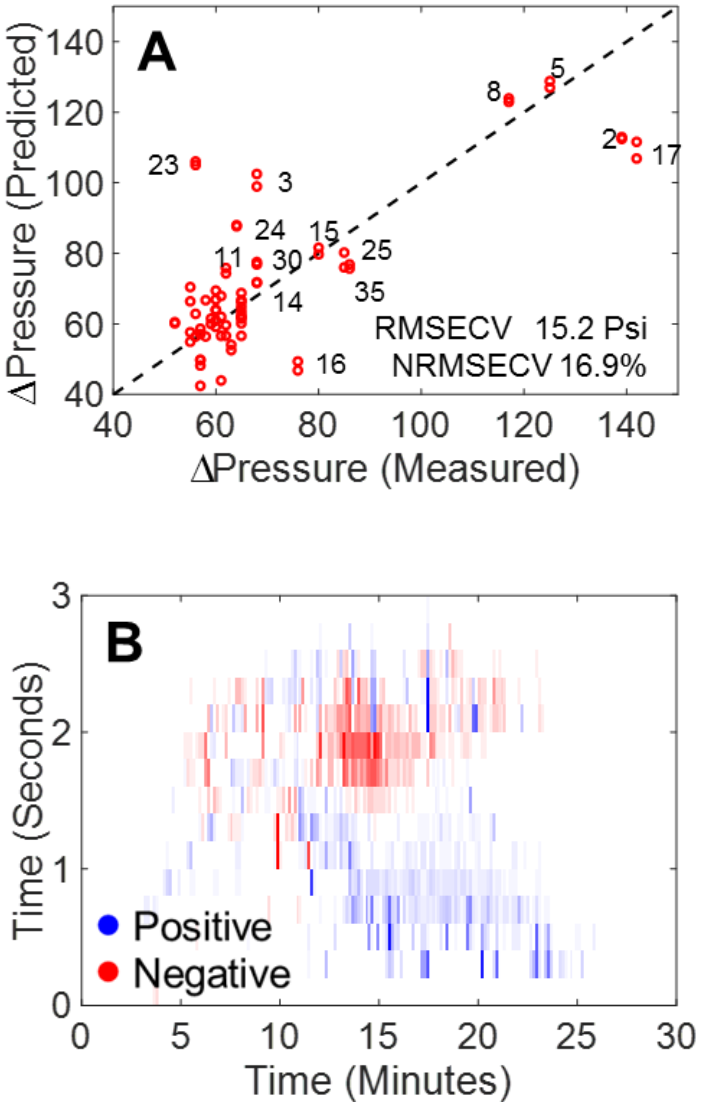


Figure C.4. Summary of the different *m/z* and their NRMSECV based on the LOOCV regressions for ΔP. (A) Mass channel is plotted as a function of NRMSECV. There does not appear to be a specific *m/z* that correlates with low NRMSECV. (B) Enhanced view of A.

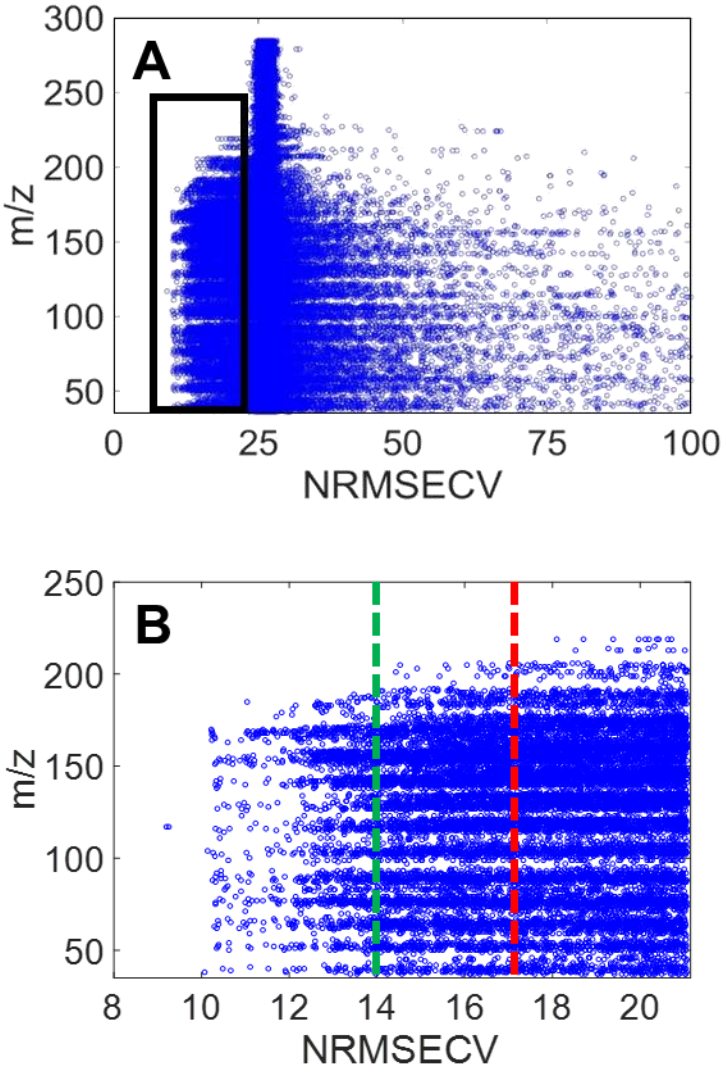


Table C.1. Summary of how the various NRMSECV thresholds impact the amount of data forwarded to PLS regression for the prediction of ΔP and the precision of the PLS model. Two thresholds of 14% and 17% were investigated.

NRMSECV threshold	# below threshold	# of tiles	RMSECV from PLS	NRMSECV from PLS
10	2	0	N/A	N/A
11	85	3	9.256	10.28
12	217	8	9.216	10.24
13	668	48	9.071	10.08
14	1677	113	9.530	10.59
15	3220	187	10.775	11.97
16	5097	272	11.119	12.35
17	7239	354	10.645	11.83
18	9418	430	11.052	12.28

Table C.2. Summary of how the various NRMSECV thresholds from the LOOCV regressions of the logarithmically scaled carbon in the heated zone impact the amount of data forwarded to PLS regression and the precision of the PLS model. Two thresholds of 17% and 18% were investigated.

NRMSECV threshold	# below threshold	# of tiles	RMSECV from PLS	NRMSECV from PLS
13	11	2	0.1887	12.54
14	65	5	0.1918	12.75
15	227	9	0.2056	13.66
16	568	24	0.2361	15.69
17	2055	120	0.2460	16.35
18	7456	422	0.2467	16.39

Table C.3. Summary of how the various NRMSECV thresholds for the LOOCV regressions of the amorphous carbon in the exit zone impact the amount of data forwarded to PLS regression and the precision of the PLS model.

NRMSECV threshold	# below threshold	# of tiles	RMSECV from PLS	NRMSECV from PLS
14	19	2	26883	13.40
15	105	11	28105	14.01
16	583	32	27630	13.77
17	1726	86	28844	14.38
18	3596	133	28899	14.41
19	6593	240	29989	14.95
20	10456	829	30649	15.28
21	14331	1049	31530	15.72
22	18029	1186	31669	15.79

Figure C.5. (A)The LOOCV regressions for the amorphous carbon in the exit zone from each tile per *m/z* which had a NRMSECV below 17% (RMSECV =36,000) and contained at least 3 *m/z* were used to create an analytical ion current chromatogram using Sample 8 to highlight the tiles identified. A total of 86 tiles were identified located predominantly within the aromatic portion of the chromatogram. (B) PLS prediction of the amorphous carbon in the exit zone was performed using the tile locations and their respective *m/z* identified in A. (C) PLS prediction of the ΔP was performed using the tile locations and their respective *m/z* identified in A.

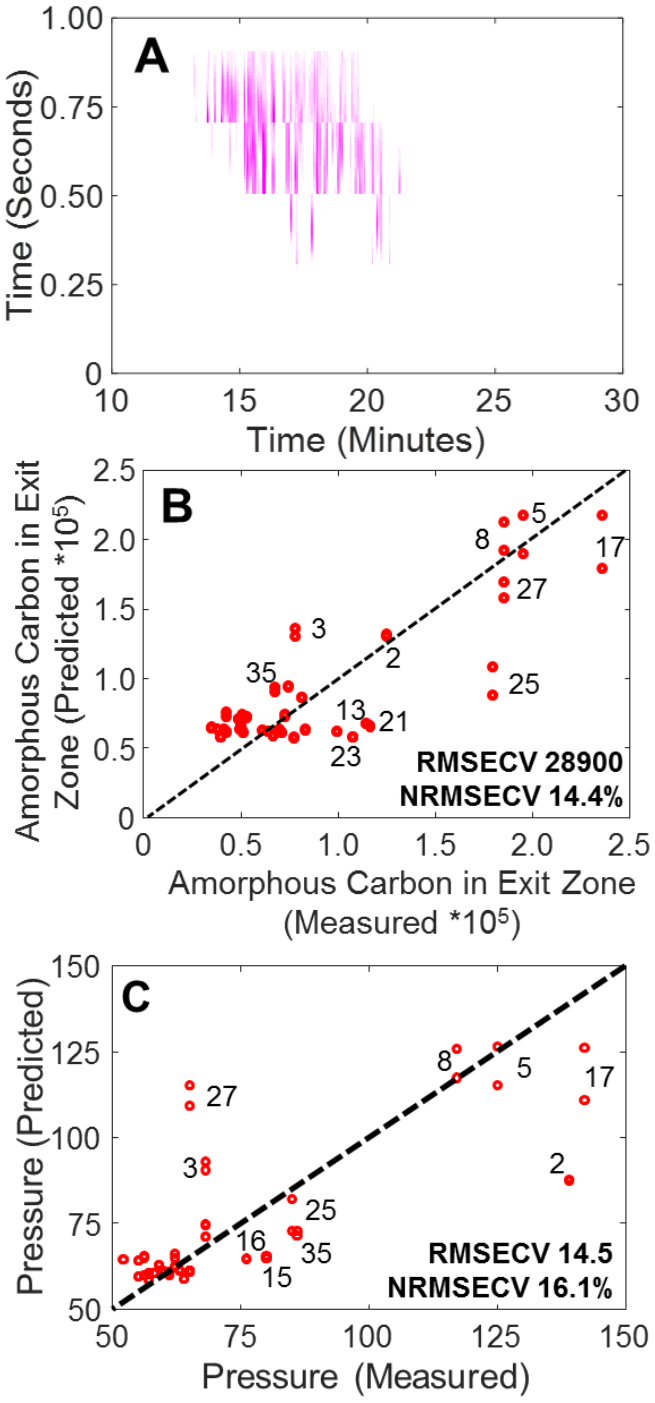


Table C.4. Summary of the top 100 tiles identified by ΔP LOOCV regressions in order of the number *m/z* identified by the LOOCV regressions of the ΔP. The retention times included are for the center of each tile.

Tile Hit Number	^2t_R (s)	^1t_R (min)	^{2}D Tile	^{1}D Tile	LRV Sign for ΔP	ΔP *m/z*	CCH *m/z*	CCE *m/z*	ACH *m/z*	ACE *m/z*	FCH *m/z*	FCE *m/z*
1	0.20	25.3	9	169	+	83	93	0	0	0	0	0
2	0.20	23.5	9	157	+	74	88	0	0	0	0	0
3	0.20	25.6	9	171	+	63	71	0	0	0	0	0
4	0.20	22.8	9	152	+	59	23	0	0	0	0	0
5	0.60	22.0	11	147	+	56	10	0	0	0	0	0
6	0.20	24.3	9	162	-	51	50	0	0	0	0	0
7	0.40	25.2	10	168	+	51	0	0	6	0	0	0
8	0.20	24.1	9	161	-	50	49	0	0	0	0	0
9	0.20	27.6	9	184	+	48	57	0	0	0	0	0
10	0.20	25.2	9	168	+	46	67	3	0	0	0	0
11	0.20	23.4	9	156	+	42	75	0	0	0	0	0
12	0.40	25.9	10	173	+	40	0	0	0	0	0	0
13	0.20	27.9	9	186	+	38	42	0	0	0	0	0
14	0.20	21.0	9	140	-	33	33	0	0	0	0	0
15	0.20	26.4	9	176	+	32	39	0	0	0	0	0
16	0.60	22.9	11	153	+	32	4	0	0	0	0	0
17	0.60	20.5	11	137	+	27	36	4	0	14	0	0
18	0.20	26.2	9	175	+	26	3	0	0	5	0	0
19	0.40	23.8	10	159	+	22	7	0	0	3	0	0
20	0.40	25.6	10	171	+	21	0	0	0	0	0	0
21	0.20	28.5	9	190	+	20	0	0	0	0	0	0
22	0.60	23.5	11	157	+	20	0	0	0	0	0	0
23	0.80	22.9	12	153	+	19	3	0	0	0	0	0
24	0.80	20.2	12	135	-	17	6	0	0	3	0	0
25	0.20	26.5	9	177	+	15	53	0	0	0	0	0
26	0.80	14.5	12	97	+	15	7	0	0	39	0	0
27	0.20	20.4	9	136	-	14	0	0	0	0	0	0
28	0.60	19.9	11	133	+	14	13	0	0	25	0	0
29	0.20	22.9	9	153	+	13	65	0	0	0	0	0
30	0.20	25.5	9	170	+	12	30	0	0	0	0	0
31	0.20	27.3	9	182	+	11	17	0	0	0	0	0
32	0.40	24.3	10	162	+	11	9	0	0	0	0	0
33	0.40	25.3	10	169	+	11	0	0	3	0	0	0
34	0.40	26.2	10	175	+	11	0	0	3	0	0	0
35	1.00	22.9	13	153	+	11	6	0	0	0	0	0
36	0.40	26.4	10	176	+	10	0	0	0	0	3	0

37	0.20	28.6	9	191	+	9	6	0	3	0	0	0
38	0.40	24.9	10	166	+	9	8	0	3	0	0	0
39	0.60	22.2	11	148	+	9	30	0	0	0	0	0
40	0.60	25.3	11	169	+	9	12	0	0	0	0	0
41	0.80	17.1	12	114	-	9	0	0	0	26	0	0
42	0.20	24.7	9	165	+	8	3	0	0	0	0	0
43	0.20	25.9	9	173	+	8	54	0	7	0	0	0
44	0.20	27.4	9	183	+	8	22	0	0	0	0	0
45	0.40	23.2	10	155	-	8	8	0	5	0	0	0
46	0.60	25.9	11	173	+	8	0	0	3	0	0	0
47	1.00	23.5	13	157	+	8	0	0	0	0	0	0
48	1.20	23.5	14	157	+	8	0	0	0	0	0	0
49	0.00	23.7	8	158	+	7	0	0	0	0	0	0
50	0.20	23.2	9	155	+	7	13	0	0	0	0	0
51	0.80	23.5	12	157	+	7	0	0	0	0	0	0
52	0.80	25.9	12	173	+	7	5	0	0	0	0	0
53	1.60	22.9	1	153	+	6	6	0	0	0	0	0
54	1.60	23.7	1	158	+	6	0	0	0	0	0	0
55	1.80	23.7	2	158	+	6	0	0	0	0	0	0
56	0.20	25.8	9	172	-	6	45	0	0	0	0	0
57	0.20	26.1	9	174	+	6	16	0	0	0	0	0
58	0.40	22.9	10	153	+	6	3	0	0	0	0	0
59	0.40	27.3	10	182	+	6	0	0	0	0	0	0
60	0.60	20.4	11	136	+	6	5	0	0	3	0	0
61	0.60	23.8	11	159	-	6	0	0	0	0	0	0
62	0.80	16.8	12	112	+	6	0	0	0	15	0	8
63	1.00	25.3	13	169	-	6	6	0	0	0	0	0
64	1.00	25.9	13	173	+	6	6	0	0	0	0	0
65	1.20	22.9	14	153	+	6	7	0	0	0	0	0
66	1.40	23.5	15	157	+	6	0	0	0	0	0	0
67	1.80	22.9	2	153	+	5	3	0	0	0	0	0
68	0.20	20.2	9	135	+	5	73	0	0	0	0	0
69	0.20	24.4	9	163	-	5	3	0	0	0	0	0
70	0.20	26.7	9	178	-	5	4	0	0	0	0	0
71	0.40	24.1	10	161	+	5	4	0	0	0	0	0
72	0.80	25.3	12	169	+	5	6	0	0	0	0	0
73	1.40	22.9	15	153	+	5	5	0	0	0	0	0
74	2.00	22.9	3	153	+	4	5	0	0	0	0	0
75	2.20	23.7	4	158	+	4	0	0	0	0	0	0
76	2.60	23.7	6	158	-	4	0	0	0	0	0	0
77	0.00	25.3	8	169	+	4	5	0	0	0	0	0
78	0.20	21.1	9	141	+	4	0	0	0	0	0	0
79	0.20	27.7	9	185	-	4	6	0	0	0	0	0
80	0.60	18.1	11	121	+	4	0	3	0	71	0	0
81	0.60	20.2	11	135	+	4	7	9	0	41	0	21
82	0.60	22.5	11	150	-	4	4	0	0	0	0	0
83	0.60	25.2	11	168	+	4	3	0	0	0	0	0
84	0.60	27.6	11	184	+	4	4	0	0	0	0	0

85	0.80	17.4	12	116	+	4	0	0	0	27	0	0
86	0.80	25.2	12	168	+	4	4	0	0	0	0	0
87	1.00	20.2	13	135	+	4	0	0	0	0	0	0
88	1.00	23.4	13	156	-	4	3	0	0	0	0	0
89	1.00	25.2	13	168	-	4	4	0	0	0	0	0
90	1.20	25.3	14	169	+	4	4	0	0	0	0	0
91	1.60	23.5	1	157	+	3	7	0	0	0	0	0
92	2.00	23.7	3	158	+	3	0	0	0	0	0	0
93	2.20	22.9	4	153	-	3	0	0	0	0	0	0
94	2.40	23.7	5	158	-	3	0	0	0	0	0	0
95	2.80	20.2	7	135	+	3	0	3	0	0	0	0
96	2.80	23.7	7	158	+	3	0	0	0	0	0	0
97	0.40	23.5	10	157	+	3	7	3	0	0	0	3
98	0.40	24.4	10	163	+	3	0	0	4	0	0	0
99	0.60	19.6	11	131	+	3	0	0	0	9	0	7
100	0.60	20.7	11	138	-	3	20	0	0	0	0	0

Appendix D

This Appendix is reproduced from the Electronic Supplementary Content of C.E. Freye, N.R. Moore, R.E. Synovec, "Enhancing the Chemical Selectivity in Discovery-based Analysis with Tandem Ionization Time-of-Flight Mass Spectrometry Detection for Comprehensive Two-Dimensional Gas Chromatography" *Journal of Chromatography A* 1537 (2018), 99-108.

Figure D.1. Visual representation of tandem ionization GC × GC – TOFMS data collection. (A) Tandem ionization allows for two different ionization energies (70 and 14 eV in this study) to be collected concurrently. The ion source switches back and forth between the two ionization energies which is represented by the green and red dots, at a rate of 50 Hz for each ionization energy. (B) The different ionization energies while collected concurrently are stored in different data files enabling the analyst to explore a single ionization energy or both simultaneously. (C) A "head-to-toe" plot of the mass spectra for the two different ionization energies. Tandem ionization has been introduced as way to obtain both hard and soft ionization for increased peak identification without decreasing sample throughput.

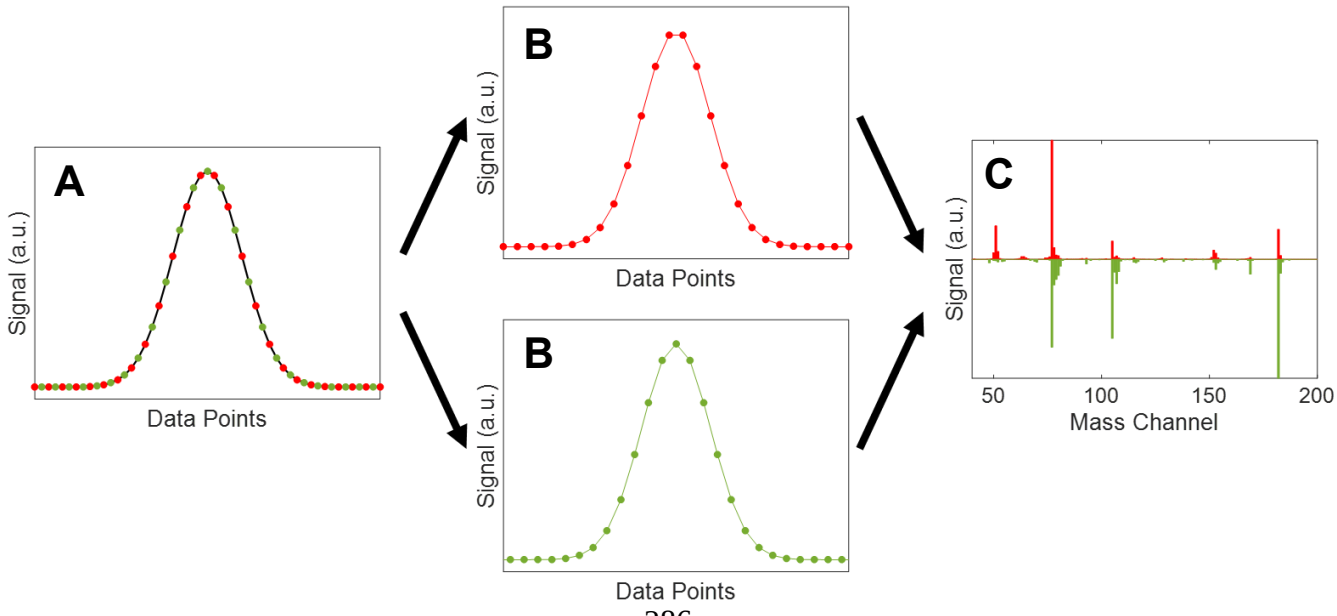


Table D.1. Table of the spiked components with the analyte number as shown in manuscript Figure 9.1. The supplier, purity, the mass in stock solution, and the actual concentrations are shown. Two analytes, t-butylbenzene and 1-ethylnapthalene are native to this diesel fuel, at 4.4 ppm and 38.2 ppm, respectively [9].

Compound Number	Name	Supplier	Purity	Mass in stock (mg)	Concentration (ppm)
1	1-Chlorohexane	Sigma-Aldrich	99%	44.7	50.0
2	2-Heptanol	Sigma-Aldrich	99%	47.8	53.4
3	Cyclooctane	Sigma-Aldrich	99%	42.8	47.9
4	2,5-Dimethylthiophene	Sigma-Aldrich	98%	37.3	41.7
5	3-Octanone	Sigma-Aldrich	98%	40.2	45.0
6	t-Butylbenzene	Sigma-Aldrich	99%	38.9	43.5
7	Limonene	Sigma-Aldrich	96%	59.7	66.8
8	5-Decyne	Sigma-Aldrich	98%	43.2	48.3
9	2-Decanone	Sigma-Aldrich	98%	44.8	50.1
10	1-Nonanol	Sigma-Aldrich	98%	56.8	63.5
11	Butyrophenone	Sigma-Aldrich	99%	41.8	46.7
12	1-Ethylnaphthalene	Sigma-Aldrich	97%	36.6	40.9

Figure D.2. Data analysis workflow for the calculation of the Fisher ratios and subsequent deconvolution by PARAFAC. Discovery of a spiked analyte is not guaranteed as it may be overlapped by a large interferent and/or be below the *S/N* filter. In addition, deconvolution by PARAFAC may not result in a successful identification due to a lack of selective *m/z* and/or large interferent.

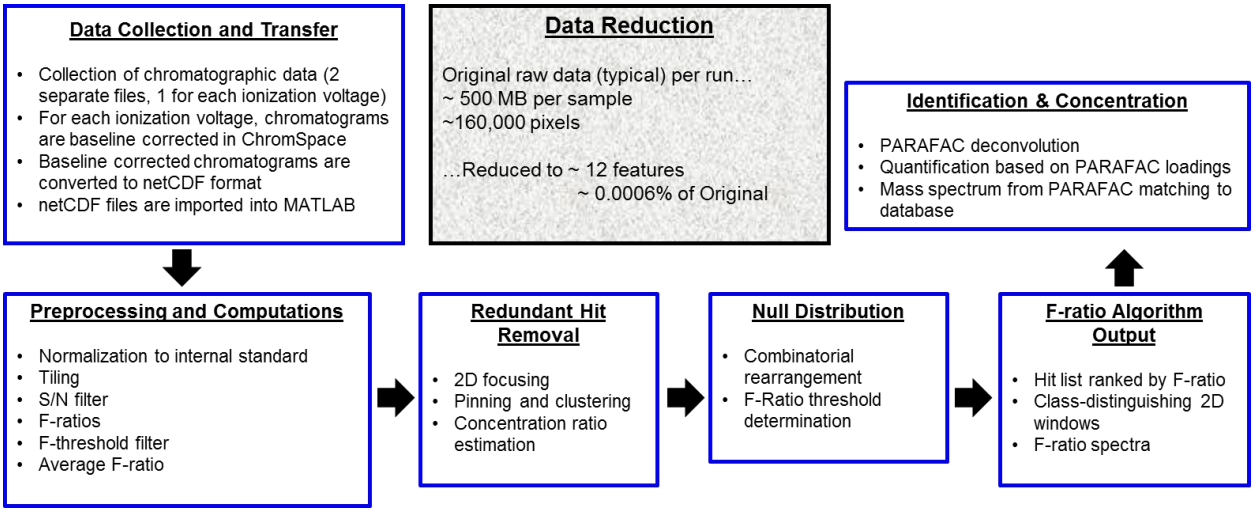


Figure D.3. (A) Total ion current (TIC) chromatogram of the 70 eV data. (B) TIC chromatogram of the 14 eV data. While the 14 eV and 70 eV TIC chromatograms look nominally the same, key differences are found at each *m/z* which is shown in the F-ratio analyses (Tables 9.1,9.2,9.3), and during the PARAFAC deconvolution providing the mass spectra.

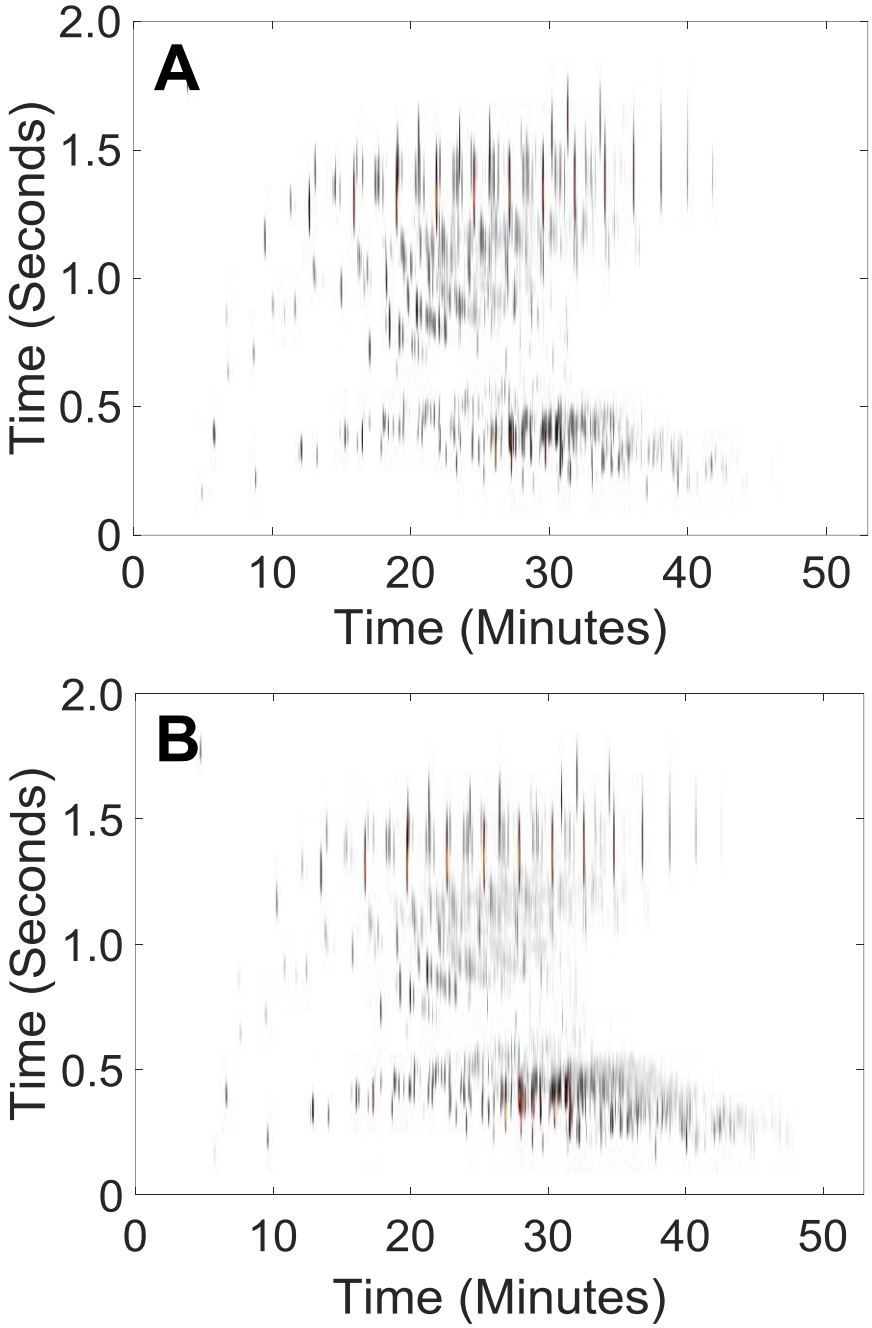
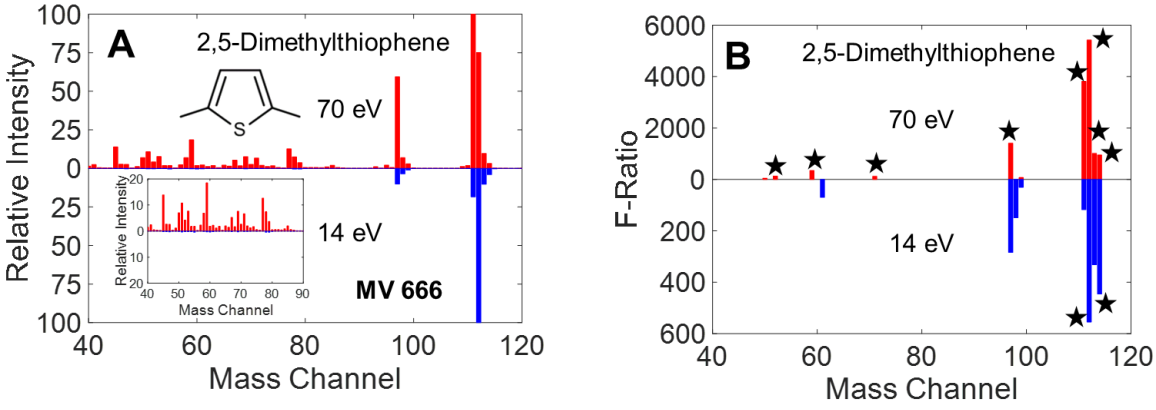
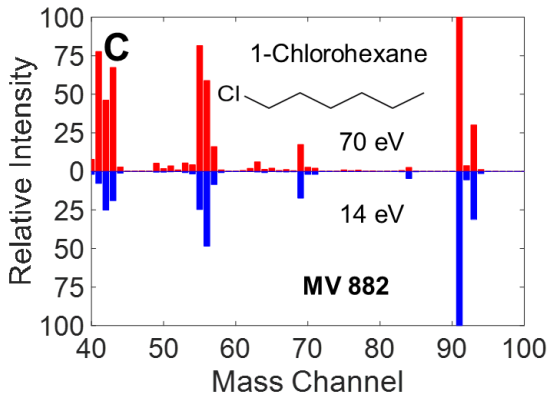
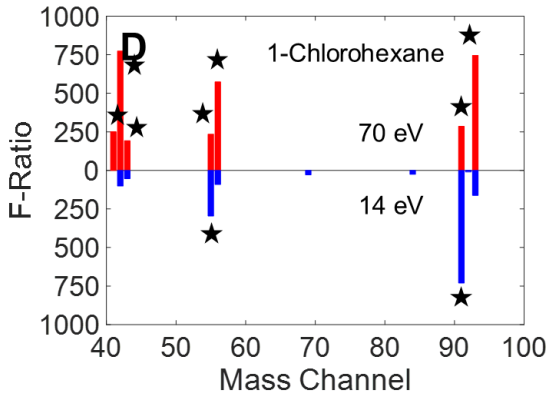
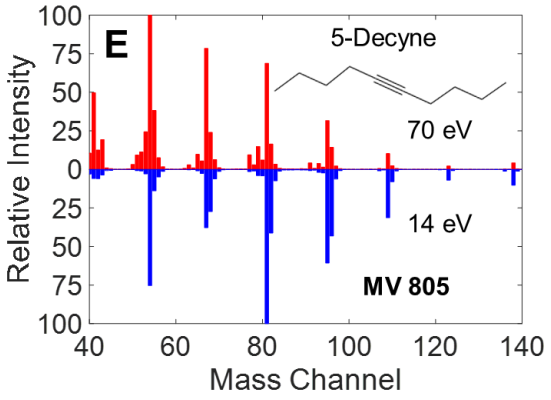


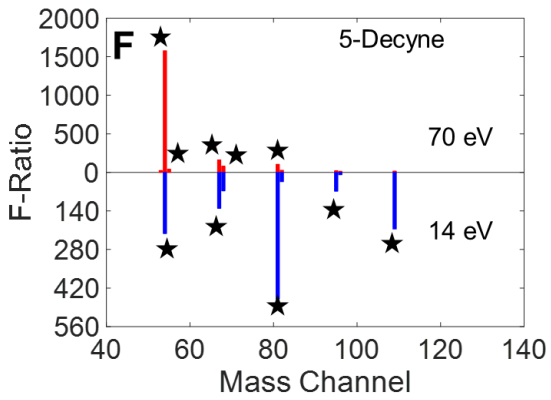
Figure D.4. Head-to-toe plots comparing the 70 eV to 14 eV mass spectra and F-ratio spectra for the remaining eleven of twelve analytes. As in Fig. 9.4 in the manuscript, the match value (MV) between the 70 and 14 eV spectra is indicated, providing a quantitative metric to infer complementariness of the mass spectra at the two ionization energies The mass channels, *m/z*, starred (★) in the F-ratio spectra are those found when the two data sets (70 eV and 14 eV) are fused and F-ratio analysis is performed. 2,5-Dimethylthiophene (A) mass spectrum, and (B) F-ratio spectrum. 1-Chlorohexane (C) mass spectrum, and (D) F-ratio spectrum. 5-Decyne (A) mass spectrum, and (F) F-ratio spectrum. t-Butylbenzene (G) mass spectrum, and (H) F-ratio spectrum. Limonene (I) mass spectrum, and (J) F-ratio spectrum. Cyclooctane (K) mass spectrum, and (L) F-ratio spectrum. 2-Decanone (M) mass spectrum, and (N) F-ratio spectrum. 3-Ocatanone (O) mass spectrum, and (P) F-ratio spectrum. 1-Nonanol (Q) mass spectrum, and (R) F-ratio spectrum. 1-Ethylnapthalene (S) mass spectrum, and (T) F-ratio spectrum. 2-Heptanol (U) mass spectrum, and (V) F-ratio spectrum.

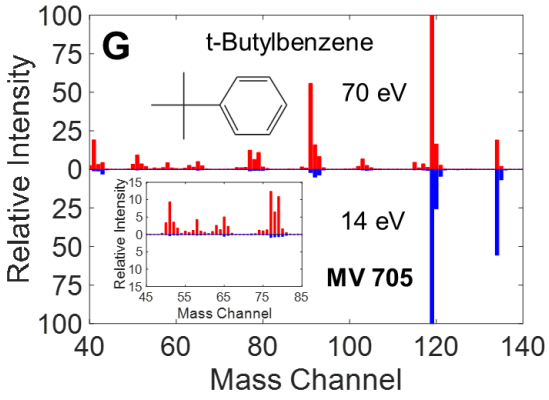


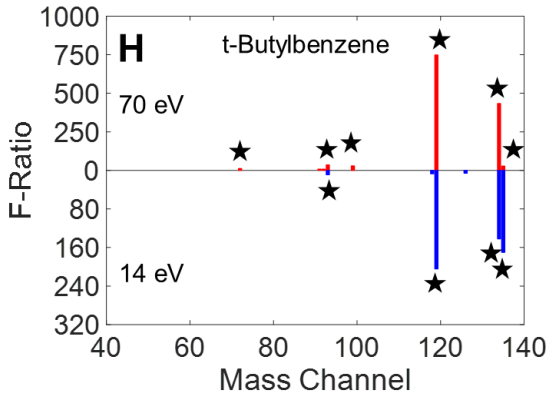

C
1-Chlorohexane
Cl
70 eV
14 eV
MV 882
Relative Intensity
Mass Channel

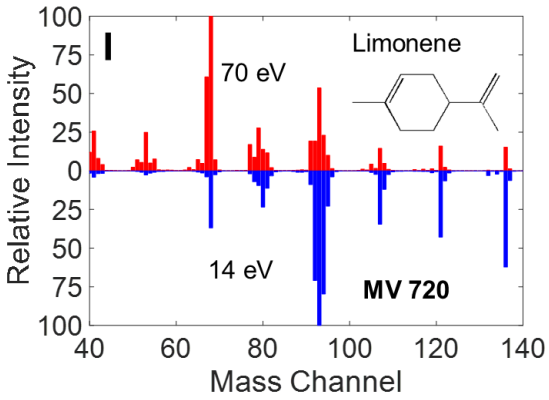

D
1-Chlorohexane
70 eV
14 eV
F-Ratio
Mass Channel

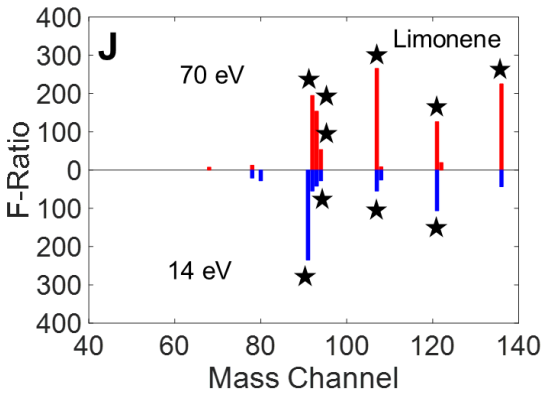

E
5-Decyne
70 eV
14 eV
MV 805
Relative Intensity
Mass Channel

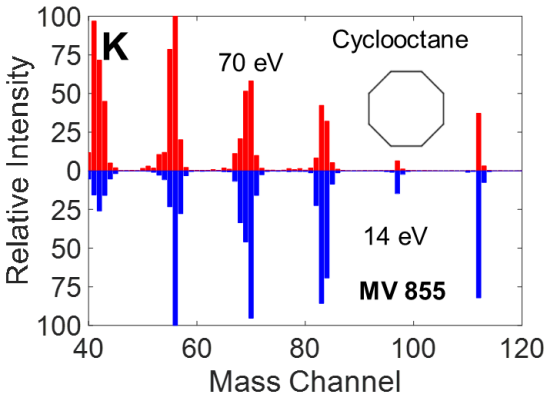

F
5-Decyne
70 eV
14 eV
F-Ratio
Mass Channel

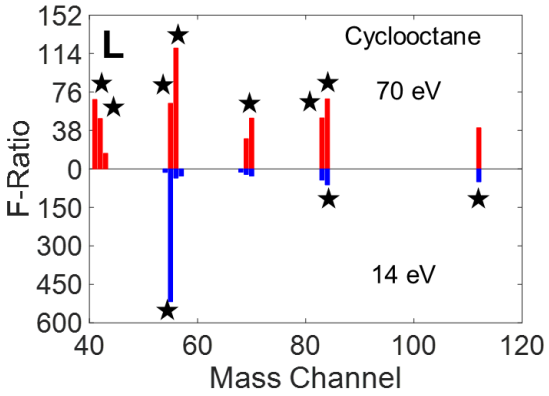

G
t-Butylbenzene
70 eV
14 eV
MV 705
Relative Intensity
Mass Channel

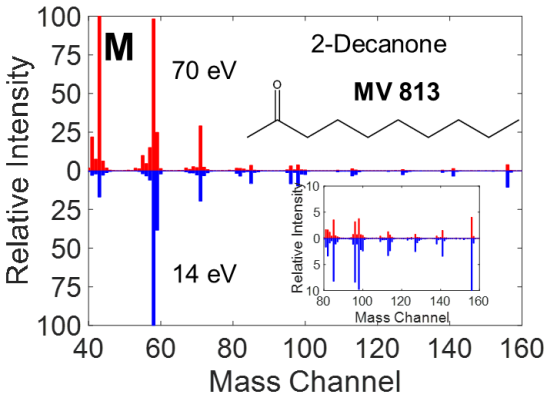

H
t-Butylbenzene
70 eV
14 eV
F-Ratio
Mass Channel

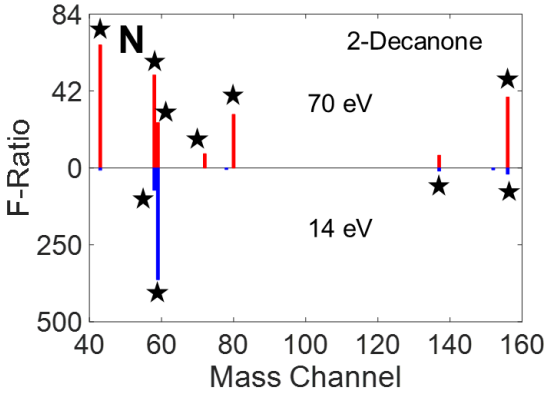

I
Limonene
70 eV
14 eV
MV 720
Relative Intensity
Mass Channel

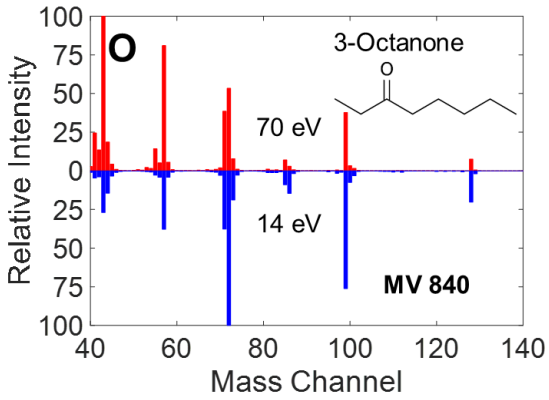

J
Limonene
70 eV
14 eV
F-Ratio
Mass Channel

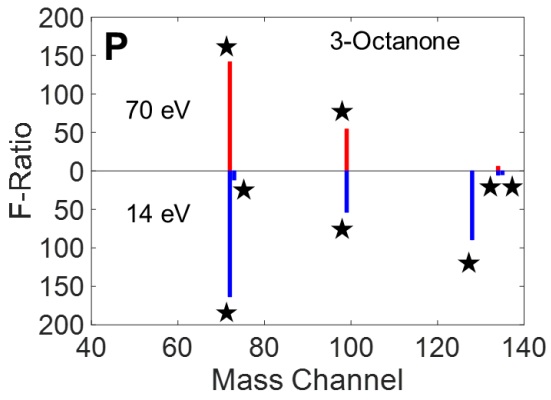

K
Cyclooctane
70 eV
14 eV
MV 855
Relative Intensity
Mass Channel

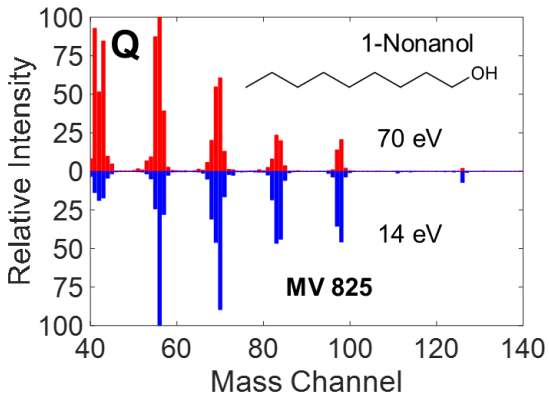

L
Cyclooctane
70 eV
14 eV
F-Ratio
Mass Channel

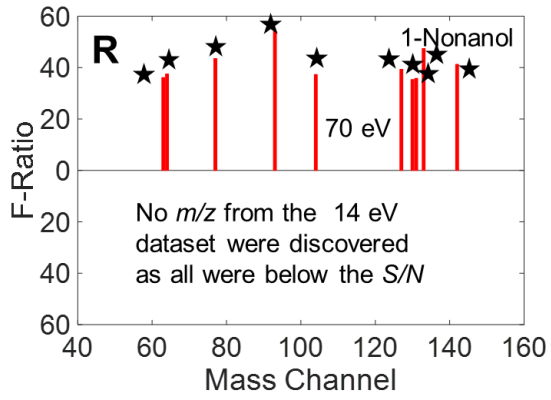

M
2-Decanone
MV 813
70 eV
14 eV
Relative Intensity
Mass Channel

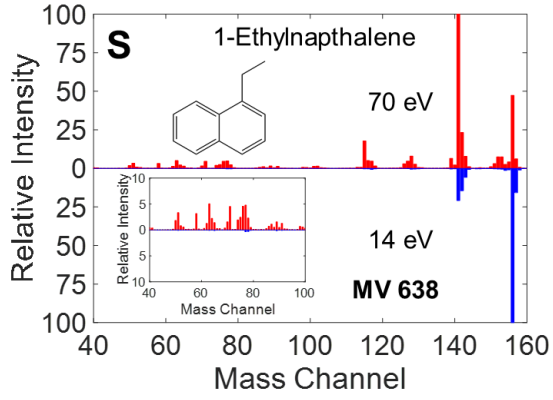

N
2-Decanone
70 eV
14 eV
F-Ratio
Mass Channel

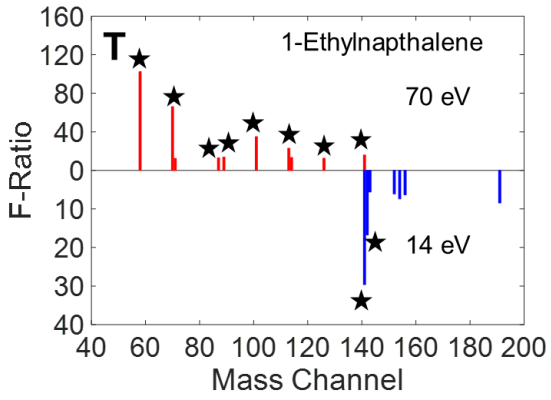

O
3-Octanone
70 eV
14 eV
MV 840
Relative Intensity
Mass Channel

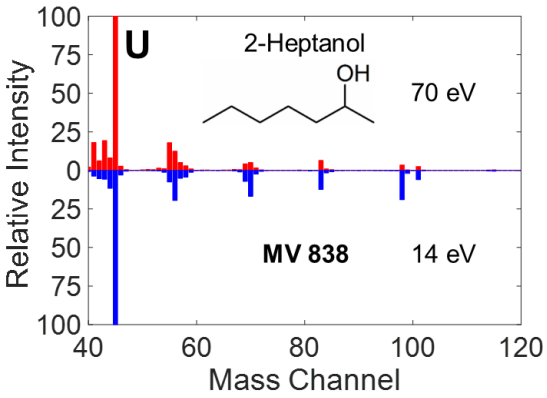

P
3-Octanone
70 eV
14 eV
F-Ratio
Mass Channel


Q
1-Nonanol
OH
70 eV
14 eV
MV 825
Relative Intensity
Mass Channel


R
1-Nonanol
70 eV
No m/z from the 14 eV
dataset were discovered
as all were below the S/N
F-Ratio
Mass Channel


S
1-Ethylnapthalene
70 eV
14 eV
MV 638
Relative Intensity
Mass Channel


T
1-Ethylnapthalene
70 eV
14 eV
F-Ratio
Mass Channel


U
2-Heptanol
OH
70 eV
MV 838
14 eV
Relative Intensity
Mass Channel
100
75
50
25
0
25
50
75
100
40
60
80
100
120

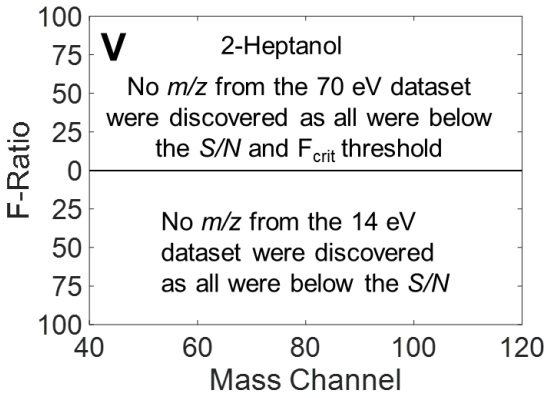

V
2-Heptanol
No m/z from the 70 eV dataset were discovered as all were below the S/N and F crit threshold
No m/z from the 14 eV dataset were discovered as all were below the S/N
F-Ratio
Mass Channel
100
75
50
25
0
25
50
75
100
40
60
80
100
120

Figure D.5. PARAFAC deconvolution of 5-Decyne at 14 eV. 5-Decyne is overlapped by two analytes. (A) ^{2}D profile which contains the relative concentration information (B) ^{1}D profile (C) Mass spectrum and match value (MV). In this instance, the MV is between the 14 eV spectrum deconvoluted by PARAFAC relative to the 14 eV library spectrum.

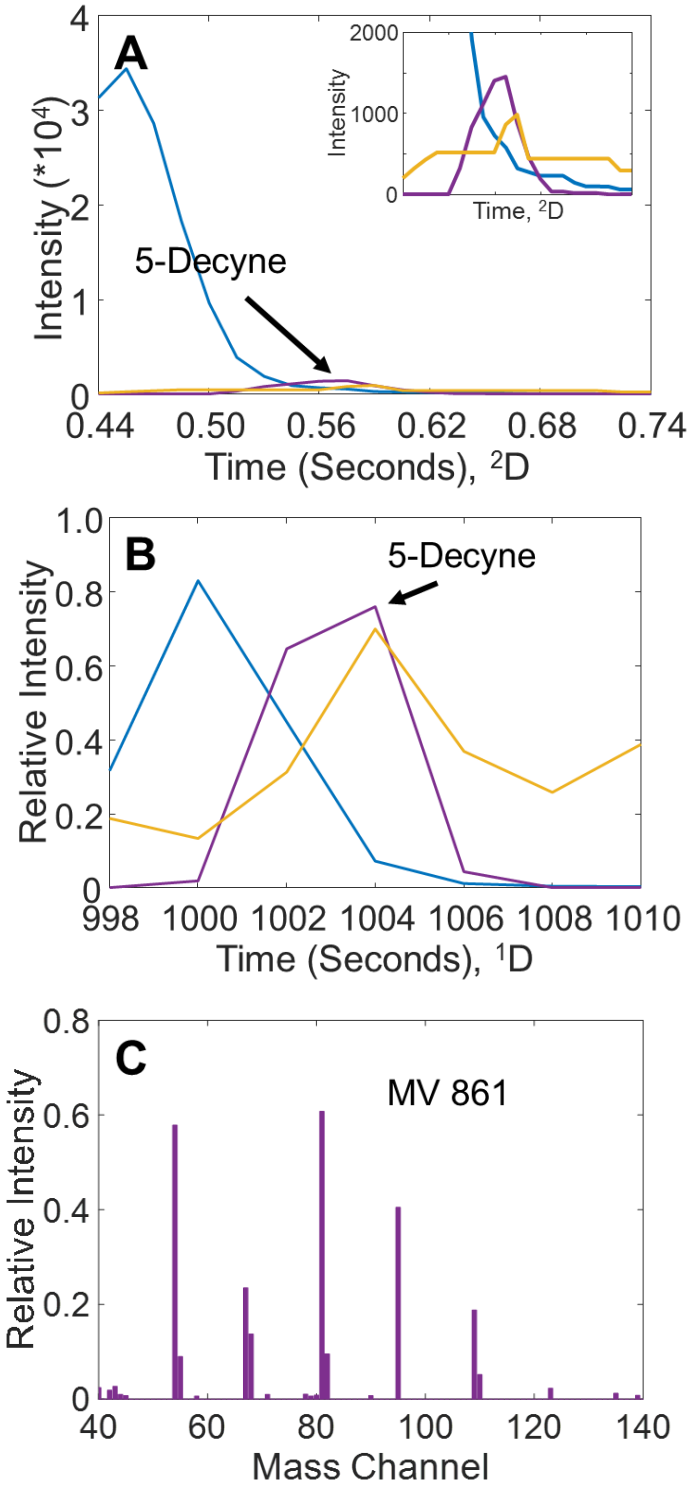
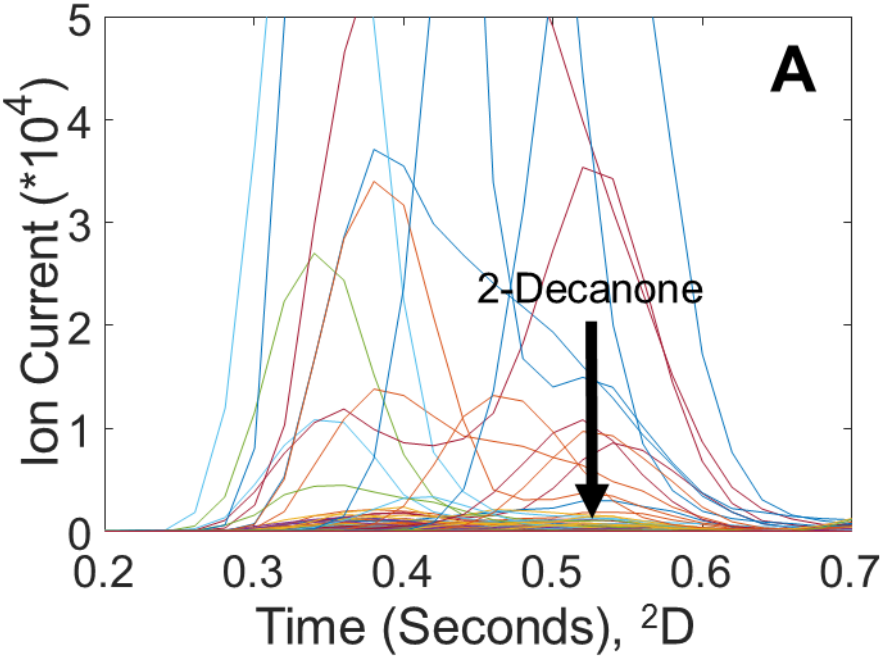
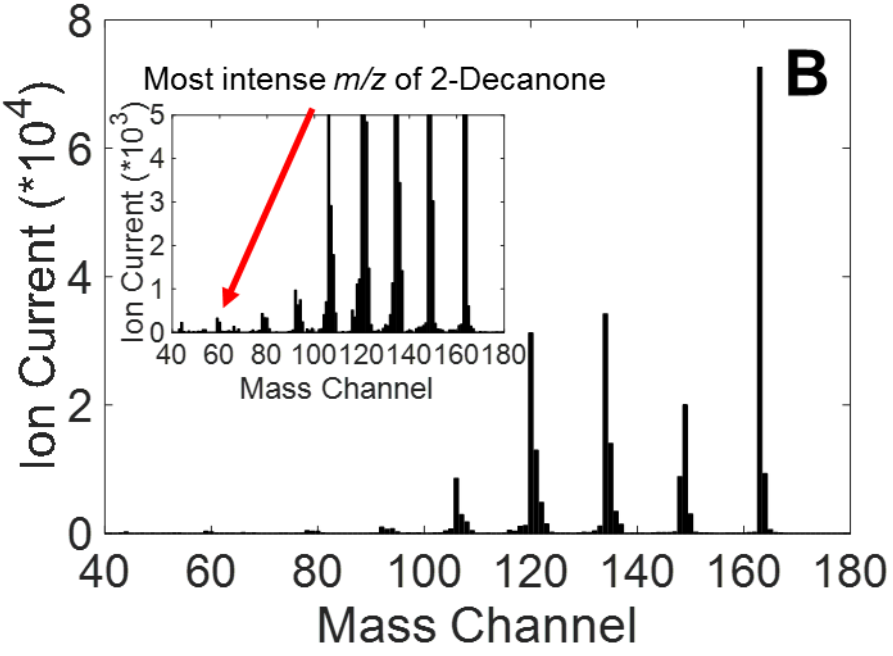


Figure D.6. (A) From the 14 eV data set, a portion of a single modulation is shown which corresponds to the highest intensity for 2-decanone which elutes at a $^{2}t_{R}$ of 0.52 s. Visually, it appears that at least four interfering analytes overlap. (B) Mass spectrum from the elution time of 2-decanone. The most intense mass channel (*m/z* 58) is ~ 10-fold less intense than the most intense *m/z*. While F-ratio analysis is able to discover the location of 2-decanone, PARAFAC deconvolution of the full mass spectrum was not successful.





Appendix E

Figure E.l. (A) This shows a zoomed view of the 2nd modulation which has a ^1t_r of 6 seconds and is centered at 60 data points in Figure 10.6. After integration, this peak is now above the LOQ and can be included for quantitation. (B**)** This shows a zoomed view of the 3rd modulation which has a ^1t_r of 9 seconds and is centered at 100 data points in Figure 10.6. This peak was originally above the LOQ but is now significantly above the LOQ.

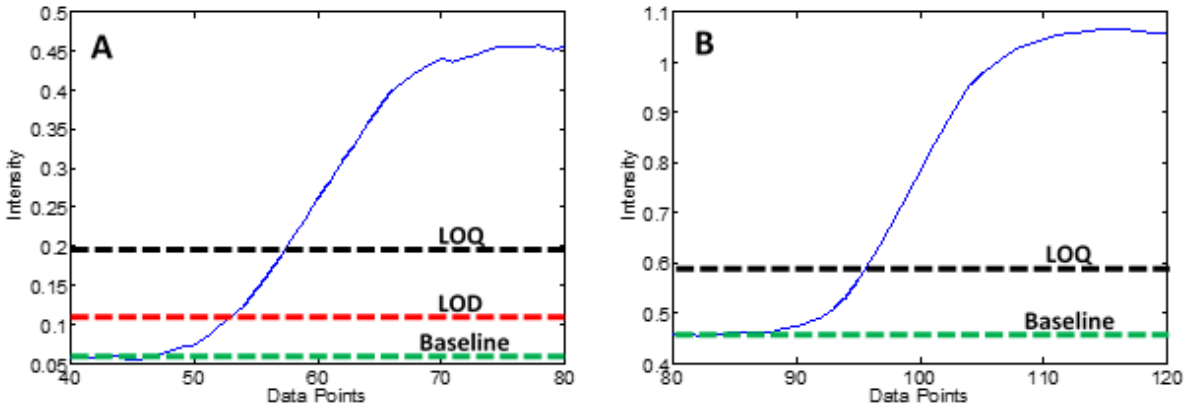
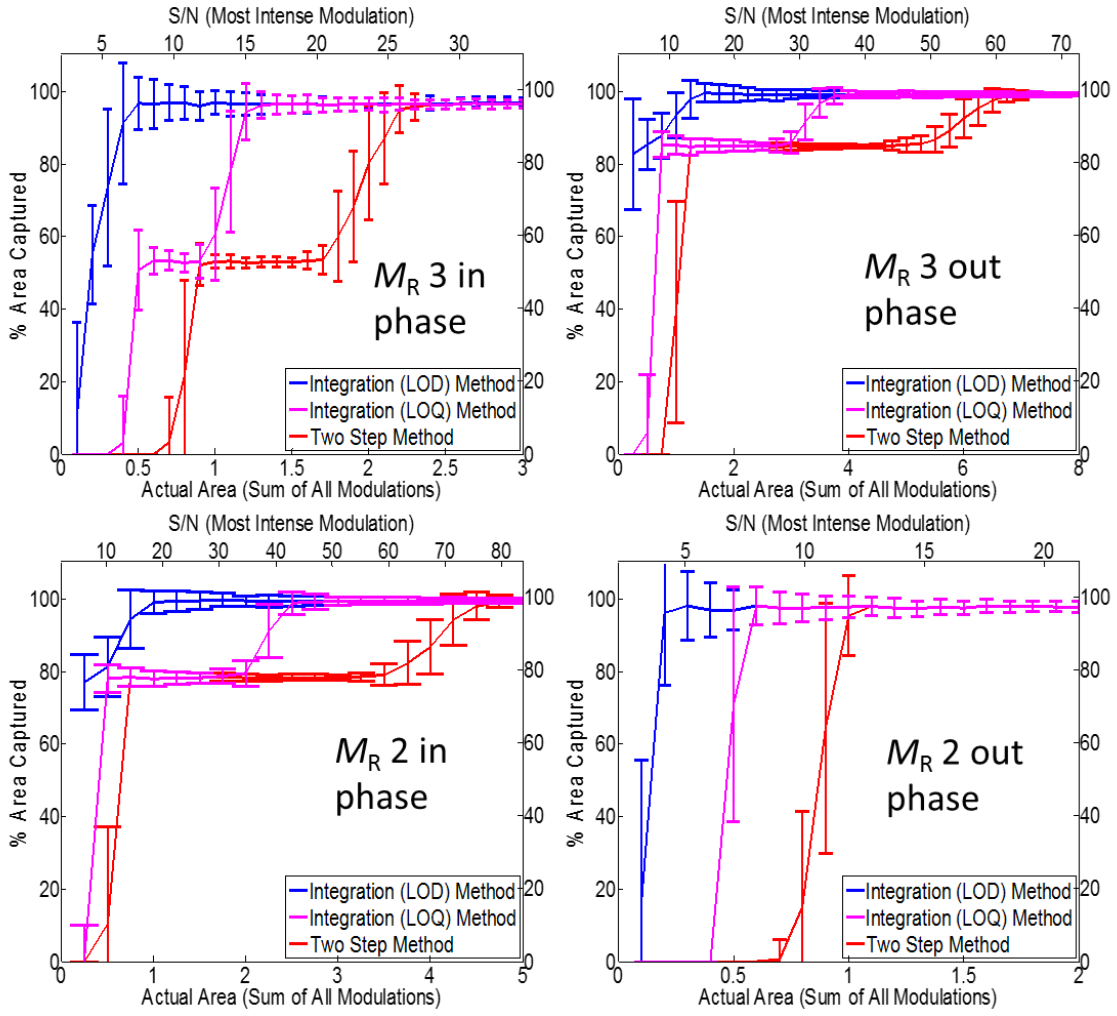


Figure E.2. Shown is the difference between the two step method and the integration method using both LOD and LOQ to determine concentration of analytes in phase and out of phase for various $M_{R.}$ (A) Shown is the comparison for a in phase peak at a M_R of 3 (B**)** Shown is the comparison for an out of phase peak at a M_R of 3 (C**)** Shown is the comparison for a in phase peak at a M_R of 2 (D**)** Shown is the comparison for an out of phase peak at a M_R of 2



BIBLIOGRAPHY

V. Abrahamsson, N. Ristic, K. Franz, K. Van Geem, Comprehensive two-dimensional gas chromatography in combination with pixel-based analysis for fouling tendency prediction, J. Chromatogr. A. 1501 (2017) 89–98.

M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC×GC): I. Introduction and instrumental set-up, TrAC. 25 (2006) 438–454.

M.S. Alam, C. Stark, R.M. Harrison, Using Variable Ionization Energy Time-of-Flight Mass Spectrometry with Comprehensive GC×GC To Identify Isomeric Species, Anal. Chem. 88 (2016) 4211–4220.

O. Altin, S. Eser, Analysis of Solid Deposits from Thermal Stressing of a JP-8 Fuel on Different Tube Surfaces in a Flow Reactor, Ind. Eng. Chem. Res. 40 (2001) 596–603.

O. Altin, S. Eser, Carbon deposit formation from thermal stressing of petroleum fuels, Am. Chem. Soc. Div. Fuel Chem. 49 (2004).

C.M. Andersen, R. Bro, Variable selection in regression—a tutorial, J. Chemometrics. 24 (2010) 728–737.

D.A. Armbruster, M.D. Tillman, L.M. Hubbs, Limit of detection (LQD)/limit of quantitation (LOQ): comparison of the empirical and the statistical methods exemplified with GC-MS assays of abused drugs., Clinical Chemistry. 40 (1994) 1233–1238.

ASTM Standard D2425-04, Standard Test Method for Hydrocarbon Types in Middle Distillates by Mass Spectrometry, ASTM International, West Conshohocken, PA, 2009.

R.T.K. Baker, D.J.C. Yates, J.A. Dumesic, Filamentous Carbon Formation over Iron Surfaces, in: Coke Formation on Metal Surfaces, American Chemical Society, 1983: pp. 1–21.

D. Ballabio, V. Consonni, Classification tools in chemistry. Part 1: linear models. PLS-DA, Anal. Methods. 5 (2013) 3790–3798.

A. Ballesteros-Gómez, J. de Boer, P.E.G. Leonards, Novel Analytical Methods for Flame Retardants and Plasticizers Based on Gas Chromatography, Comprehensive Two-Dimensional Gas Chromatography, and Direct Probe Coupled to Atmospheric Pressure Chemical Ionization-High Resolution Time-of-Flight-Mass Spectrometry, Anal. Chem. 85 (2013) 9572–9580.

R. Bates, J. Edwards, M. Meyer, Heat Transfer and Deposition Behavior of Hydrocarbon Rocket Fuels, in: 41st Aerospace Sciences Meeting and Exhibit, American Institute of Aeronautics and Astronautics, n.d.

H.D. Bean, J.-M.D. Dimandja, J.E. Hill, Bacterial volatile discovery using solid phase microextraction and comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, J. Chromatogr. B. 901 (2012) 41–46.

K. Beebe, Chemometrics: A Practical Guide, Wiley, 1998.

J. Beens, M. Adahchour, R.J.J. Vreuls, K. van Altena, U.A. Th. Brinkman, Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 919 (2001) 127–132.

N.J. Begue, J.A. Cramer, C. Von Bargen, K.M. Myers, K.J. Johnson, R.E. Morris, Automated Method for Determining Hydrocarbon Distributions in Mobility Fuels, Energy Fuels. 25 (2011) 1617–1623.

E. Bernal, Limit of Detection and Limit of Quantification Determination in Gas Chromatography, (2014).

M. Billingsley, T. Edwards, L. Shafer, T. Bruno, Extent and Impacts of Hydrocarbon Fuel Compositional Variability for Aerospace Propulsion Systems, in: 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, American Institute of Aeronautics and Astronautics, n.d.

L.M. Blumberg, Accumulating resampling (modulation) in comprehensive two-dimensional capillary GC (GC×GC), J. Sep. Science. 31 (2008) 3358–3365.

A.G. Borsa, A.M. Herring, J.T. McKinnon, R.L. McCormick, G.H. Ko, Coke and Byproduct Formation during 1,2-Dichloroethane Pyrolysis in a Laboratory Tubular Reactor, Ind. Eng. Chem. Res. 40 (2001) 2428–2436.

P.W.J.M. Boumans, A tutorial review of some elementary concepts in the statistical evaluation of trace element measurements, Spectrochimica Acta Part B: Atomic Spectroscopy. 33 (1978) 625–634.

R. Bro, PARAFAC. Tutorial and applications, Chem. Intel. Lab. Sys. 38 (1997) 149–171.

C.A. Bruckner, B.J. Prazen, R.E. Synovec, Comprehensive Two-Dimensional High-Speed Gas Chromatography with Chemometric Analysis, Anal. Chem. 70 (1998) 2796–2804.

T.J. Bruno, L.S. Ott, T.M. Lovestead, M.L. Huber, The composition-explicit distillation curve technique: Relating chemical analysis and physical properties of complex fluids, , J. Chromatogr. A. 1217 (2010) 2703–2715.

P.A. Bueno, J.V. Seeley, Flow-switching device for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1027 (2004) 3–10.

B.V. Burger, T. Snyman, W.J.G. Burger, W.F. van Rooyen, Thermal modulator array for analyte modulation and comprehensive two-dimensional gas chromatography, J. Sep. Science. 26 (2003) 123–128.

J.L. Burger, T.J. Bruno, Application of the Advanced Distillation Curve Method to the Variability of Jet Fuels, Energy Fuels. 26 (2012) 3661–3671.

H. Cai, S.D. Stearns, Partial Modulation Method via Pulsed Flow Modulator for Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 76 (2004) 6064–6076.

W. Cai, Y. Li, X. Shao, A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra, Chem. Intel. Lab. Sys. 90 (2008) 188–194.

J.B. Callis, Super Resolution in Chromatography by Numerical Deconvolution, in: Ultrahigh Resolution Chromatography, American Chemical Society, 1984: pp. 171–198.

K.H. Chang, Limit of Detection and Its Establishment in Analytical Chemistry, Health and the Environmental Journal. 2 (2011) 38–43.

J. Chen, N. Mclean, D. Hager, Prediction of Molecular Weight By-Boiling-Point Distribution of Middle Distillates from Gas Chromatography−Field Ionization Mass Spectrometry (GC−FIMS), Energy Fuels. 25 (2011) 719–726.

T.-Y. Chen, M.-J. Li, J.-L. Wang, Sub-second thermal desorption of a micro-sorbent trap for the analysis of ambient volatile organic compounds, J. Chromatogr. A. 976 (2002) 39–45.

J.S. Chickos, A.E. Wentz, D. Hillesheim-Cox, M.J. Zehe, Measurement of the Vaporization Enthalpy of Complex Mixtures by Correlation-Gas Chromatography. The Vaporization Enthalpy of RJ-4, a High-Energy-Density Rocket Fuel at T = 298.15 K, Ind. Eng. Chem. Res. 42 (2003) 2874–2877.

G.D. Christian, Analytical Chemistry, 6th ed., John Wiley & Sons, Inc., Hoboken, NJ, 2004.

D.J. Cookson, B.E. Smith, Calculation of jet and diesel fuel properties using carbon-13 NMR spectroscopy, Energy Fuels. 4 (1990) 152–156.

R.B. Crosier, A New Two-Sided Cumulative Sum Quality Control Scheme, Technometrics. 28 (1986) 187–194.

L.A. Currie, The limit of precision in nuclear and analytical chemistry, Nuclear Instruments and Methods. 100 (1972) 387–395.

J. Dallüge, M. van Rijn, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection applied to the determination of pesticides in food extracts, J. Chromatogr. A. 965 (2002) 207–217.

J. Dallüge, L.L.P. van Stee, X. Xu, J. Williams, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Unravelling the composition of very complex samples by comprehensive gas chromatography coupled to time-of-flight mass spectrometry: Cigarette smoke, J. Chromatogr. A. 974 (2002) 169–184.

M.K. Das, S.C. Bishwal, A. Das, D. Dabral, A. Varshney, V.K. Badireddy, R. Nanda, Investigation of Gender-Specific Exhaled Breath Volatome in Humans by GCxGC-TOF-MS, Anal. Chem. 86 (2014) 1229–1237.

J.M. Davis, D.R. Stoll, P.W. Carr, Effect of First-Dimension Undersampling on Effective Peak Capacity in Comprehensive Two-Dimensional Separations, Anal. Chem. 80 (2008) 461–473.

D.R. Deans, A new technique for heart cutting in gas chromatography, Chromatographia. 1 (1968) 18–22.

M.M. van Deursen, J. Beens, H.-G. Janssen, P.A. Leclercq, C.A. Cramers, Evaluation of time-of-flight mass spectrometric detection for fast gas chromatography, J. Chromatogr. A. 878 (2000) 205–213.

M.J. DeWitt, T. Edwards, L. Shafer, D. Brooks, R. Striebich, S.P. Bagley, M.J. Wornat, Effect of Aviation Fuel Type on Pyrolytic Reactivity and Deposition Propensity under Supercritical Conditions, Ind. Eng. Chem. Res. 50 (2011) 10434–10451.

J.-M.D. Dimandja, S.B. Stanfill, J. Grainger, D.G. Patterson, Application of Comprehensive Two-Dimensional Gas Chromatography (GC×GC) to the Qualitative Analysis of Essential Oils, J. High Resol. Chromatogr. 23 (2000) 208–214.

S. Du, J. A. Valla, G. M. Bollas, Characteristics and origin of char and coke from fast and slow, catalytic and thermal pyrolysis of biomass and relevant model compounds, Green Chemistry. 15 (2013) 3214–3229.

L.M. Dubois, K.A. Perrault, P.-H. Stefanuto, S. Koschinski, M. Edwards, L. McGregor, J.-F. Focant, Thermal desorption comprehensive two-dimensional gas chromatography coupled to variable-energy electron ionization time-of-flight mass spectrometry for monitoring subtle changes in volatile organic compound profiles of human blood, J. Chromatogr. A. 1501 (2017) 117–127.

M. Edwards, H. Boswell, T. Górecki, Comprehensive Multidimensional Chromatography, Current Chromatography. 2 (2015) 80–109.

T. Edwards, Cracking and Deposition Behavior of Supercritical Hydrocarbon Aviation Fuels, Combust. Sci. Technol. 178 (2006) 307–334.

A. Eftekhari, H. Parastar, Multivariate analytical figures of merit as a metric for evaluation of quantitative measurements using comprehensive two-dimensional gas chromatography–mass spectrometry, J. Chromatogr. A. 1466 (2016) 155–165.

P.H. Ellaway, Cumulative sum technique and its application to the analysis of peristimulus time histograms, Electroencephalography and Clinical Neurophysiology. 45 (1978) 302–304.

S. Eser, R. Venkataraman, O. Altin, Deposition of Carbonaceous Solids on Different Substrates from Thermal Stressing of JP-8 and Jet A Fuels, Ind. Eng. Chem. Res. 45 (2006) 8946–8955.

L.S. Ettre, Development of chromatography, Anal. Chem. 43 (1971) 20A–31A.

B.M. Fabuss, J.O. Smith, C.N. Satterfield, Advances in Petroleum Chemistry Refining, Interscience, n.d.

R. Fernández-Varela, J.M. Andrade, S. Muniategui, D. Prada, F. Ramírez-Villalobos, The comparison of two heavy fuel oils in composition and weathering pattern, based on IR, GC-FID and GC–MS analyses: Application to the Prestige wreackage, Water Research. 43 (2009) 1015–1026.

A.C. Fieldner, W.A. Selvig, Use of the hydrogen-volatile-matter ratio in obtaining the net heating value of American coals, Govt. print. off., 1918.

R.A. Fisher, Statistical Methods for Research Workers, 14th ed., Oliver and Boyd, 1970.

J.P. Foley, J.G. Dorsey, Clarification of the limit of detection in chromatography, Chromatographia. 18 (1984) 503–511.

S.L. Forbes, A.N. Troobnikoff, M. Ueland, K.D. Nizio, K.A. Perrault, Profiling the decomposition odour at the grave surface before and after probing, Forensic Science International. 259 (2016) 193–199.

T.J. Fortin, Assessment of Variability in the Thermophysical Properties of Rocket Propellant RP-1, Energy & Fuels. 26 (2012) 6501–6507.

T.J. Fortin, T.J. Bruno, Assessment of the Thermophysical Properties of Thermally Stressed RP-1 and RP-2, Energy Fuels. 27 (2013) 2506–2514.

C.G. Fraga, Chemometric approach for the resolution and quantification of unresolved peaks in gas chromatography–selected-ion mass spectrometry data, J. Chromatogr. A. 1019 (2003) 31–42.

C.G. Fraga, C.A. Bruckner, R.E. Synovec, Increasing the Number of Analyzable Peaks in Comprehensive Two-Dimensional Separations through Chemometrics, Anal. Chem. 73 (2001) 675–683.

C.G. Fraga, B.J. Prazen, R.E. Synovec, Comprehensive Two-Dimensional Gas Chromatography and Chemometrics for the High-Speed Quantitative Analysis of Aromatic Isomers in a Jet Fuel Using the Standard Addition Method and an Objective Retention Time Alignment Algorithm, Anal. Chem. 72 (2000) 4154–4162.

R. French, S. Czernik, Catalytic pyrolysis of biomass for biofuels production, Fuel Processing Technology. 91 (2010) 25–32.

C.E. Freye, H.D. Bahaghighat, R.E. Synovec, Comprehensive two-dimensional gas chromatography using partial modulation via a pulsed flow valve with a short modulation period, Talanta. 177 (2018) 142–149.

C.E. Freye, B.D. Fitz, M.C. Billingsley, R.E. Synovec, Partial least squares analysis of rocket propulsion fuel data using diaphragm valve-based comprehensive two-dimensional gas chromatography coupled with flame ionization detection, Talanta. 153 (2016) 203–210.

C.E. Freye, N.R. Moore, R.E. Synovec, Enhancing the chemical selectivity in discovery-based analysis with tandem ionization time-of-flight mass spectrometry detection for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1537 (2018) 99–108.

C.E. Freye, L. Mu, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1424 (2015) 127–133.

C.E. Freye, R.E. Synovec, High temperature diaphragm valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry, Talanta. 161 (2016) 675–680.

X. Gao, L. Pang, S. Zhu, W. Zhang, W. Dai, D. Li, S. He, Gas purge microsyringe extraction coupled to comprehensive two-dimensional gas chromatography for the characterization of petroleum migration, Organic Geochemistry. 106 (2017) 30–47.

N. Gascoin, G. Abraham, P. Gillard, Synthetic and jet fuels pyrolysis for cooling and combustion applications, Journal of Analytical and Applied Pyrolysis. 89 (2010) 294–306.

P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta. 185 (1986) 1–17.

W. Genuit, H. Chaabani, Comprehensive two-dimensional gas chromatography-field ionization time-of-flight mass spectrometry (GCxGC-FI-TOFMS) for detailed hydrocarbon middle distillate analysis, Int. J. Mass Spectrom. 413 (2017) 27–32.

H.-J. de Geus, J. de Boer, U.A.T. Brinkman, Development of a thermal desorption modulator for gas chromatography, J. Chromatogr. A. 767 (1997) 137–151.

H.-J. de Geus, A. Schelvis, J. de Boer, U.A.T. Brinkman, Comprehensive Two-Dimensional Gas Chromatography with a Rotating Thermal Desorption Modulator and Independently Temperature-Programmable Columns, J. High Resol. Chromatogr. 23 (2000) 189–196.

J.C. Giddings, Two-dimensional separations: concept and promise, Anal. Chem. 56 (1984) 1258A–1270A.

J.C. Giddings, Unified separation science, 1st ed., Wiley-Interscience, 1991.

A. Giri, M. Coutriade, A. Racaud, K. Okuda, J. Dane, R.B. Cody, J.-F. Focant, Molecular Characterization of Volatiles and Petrochemical Base Oils by Photo-Ionization GC×GC-TOF-MS, Anal. Chem. 89 (2017) 5395–5403.

L.A.F. de Godoy, M.P. Pedroso, L.W. Hantao, R.J. Poppi, F. Augusto, Quantitative analysis by comprehensive two-dimensional gas chromatography using interval Multi-way Partial Least Squares calibration, Talanta. 83 (2011) 1302–1307.

R.V. Gough, T.J. Bruno, Composition-Explicit Distillation Curves of Alternative Turbine Fuels, Energy Fuels. 27 (2013) 294–302.

A.A. Gowen, G. Downey, C. Esquerre, C.P. O'Donnell, Preventing over-fitting in PLS calibration models of near-infrared (NIR) spectroscopy data using regression coefficients, J. Chemometrics. 25 (2011) 375–381.

J.F. Griffith, W.L. Winniford, K. Sun, R. Edam, J.C. Luong, A reversed-flow differential flow modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1226 (2012) 116–123.

R.L. Grob, E.F. Barry, eds., Modern Practice of Gas Chromatography, 4th ed., John Wiley & Sons, Inc., n.d.

T. Gröger, M. Schäffer, M. Pütz, B. Ahrens, K. Drew, M. Eschner, R. Zimmermann, Application of two-dimensional gas chromatography combined with pixel-based chemometric processing for the chemical profiling of illicit drug samples, J. Chromatogr. A. 1200 (2008) 8–16.

G.M. Gross, B.J. Prazen, J.W. Grate, R.E. Synovec, High-Speed Gas Chromatography Using Synchronized Dual-Valve Injection, Anal. Chem. 76 (2004) 3517–3524.

Q. Gu, F. David, F. Lynen, K. Rumpel, G. Xu, P. De Vos, P. Sandra, Analysis of bacterial fatty acids by flow modulated comprehensive two-dimensional gas chromatography with parallel flame ionization detector/mass spectrometry, J. Chromatogr. A. 1217 (2010) 4448–4453.

J.F. Hamilton, P.J. Webb, A.C. Lewis, J.R. Hopkins, S. Smith, P. Davy, Partially oxidised organic components in urban aerosol using GCXGC-TOF/MS, Atmos. Chem. Phys. 4 (2004) 1279–1290.

L. Hanley, R. Zimmermann, Light and Molecular Ions: The Emergence of Vacuum UV Single-Photon Ionization in MS, Anal. Chem. 81 (2009) 4174–4182.

P.M. Harvey, R.A. Shellie, Data Reduction in Comprehensive Two-Dimensional Gas Chromatography for Rapid and Repeatable Automated Data Analysis, Anal. Chem. 84 (2012) 6501–6507.

J. Harynuk, T. Górecki, New liquid nitrogen cryogenic modulator for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1019 (2003) 53–63.

J. Harynuk, T. Górecki, Comprehensive two-dimensional gas chromatography in stop-flow mode, J. Sep. Science. 27 (2004) 431–441.

D.M. Hawkins, D.H. Olwell, Cumulative Sum Charts and Charting for Quality Improvement, Springer Science & Business Media, 2012.

V. Heines, Chromatography-history of a parallel development, Chem. Technol. 1 (1971) 280–285.

Y.V. Heyden, R. Boque, Practical Data Handling-The Limit of Detection-The concept of the limit of detection (LOD) often arouses controversy in analytical chemistry, LCGC Europe. 23 (2009) 82–85.

J.C. Hoggard, R.E. Synovec, Parallel Factor Analysis (PARAFAC) of Target Analytes in GC × GC−TOFMS Data: Automated Selection of a Model with an Appropriate Number of Factors, Anal. Chem. 79 (2007) 1611–1619.

J.C. Hoggard, R.E. Synovec, Automated Resolution of Nontarget Analyte Signals in GC × GC-TOFMS Data Using Parallel Factor Analysis, Anal. Chem. 80 (2008) 6677–6688.

J.L. Hope, K.J. Johnson, M.A. Cavelti, B.J. Prazen, J.W. Grate, R.E. Synovec, High-speed gas chromatographic separations with diaphragm valve-based injection and chemometric analysis as a gas chromatographic "sensor," Anal. Chim. Acta. 490 (2003) 223–230.

P.Y. Hsieh, K.R. Abel, T.J. Bruno, Analysis of Marine Diesel Fuel with the Advanced Distillation Curve Method, Energy Fuels. 27 (2013) 804–810.

H. Huang, L. Spadaccini, D. Sobel, Endothermic Heat-Sink of Jet Fuels for Scramjet Cooling, in: 2002.

H. Huang, L.J. Spadaccini, D.R. Sobel, Fuel-Cooled Thermal Management for Advanced Aeroengines, J. Eng. Gas Turbines Power. 126 (2004) 284–293.

X. Huang, F.E. Regnier, Differential Metabolomics Using Stable Isotope Labeling and Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry, Anal. Chem. 80 (2008) 107–114.

M.L. Huber, E.W. Lemmon, T.J. Bruno, Effect of RP-1 Compositional Variability on Thermophysical Properties, Energy Fuels. 23 (2009) 5550–5555.

C. Hurtado, H. Parastar, V. Matamoros, B. Piña, R. Tauler, J.M. Bayona, Linking the morphological and metabolomic response of Lactuca sativa L exposed to emerging contaminants using GC × GC-MS and chemometric tools, Scientific Reports. 7 (2017).

T. Hyötyläinen, M. Kallio, K. Hartonen, M. Jussila, S. Palonen, M.-L. Riekkola, Modulator Design for Comprehensive Two-Dimensional Gas Chromatography: Quantitative Analysis of Polyaromatic Hydrocarbons and Polychlorinated Biphenyls, Anal. Chem. 74 (2002) 4441–4446.

A.T. James, A.J.P. Martin, Gas-liquid partition chromatography: the separation and micro-estimation of volatile fatty acids from formic acid to dodecanoic acid, Biochem J. 50 (1952) 679–690.

M.K. Jennerwein, M. Eschner, T. Gröger, T. Wilharm, R. Zimmermann, Complete Group-Type Quantification of Petroleum Middle Distillates Based on Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry (GC×GC-TOFMS) and Visual Basic Scripting, Energy Fuels. 28 (2014) 5670–5681.

H.M. Jeong, M.W. Seo, S.M. Jeong, B.K. Na, S.J. Yoon, J.G. Lee, W.J. Lee, Pyrolysis kinetics of coking coal mixed with biomass under non-isothermal and isothermal conditions, Bioresource Technology. 155 (2014) 442–445.

R. Jiang, G. Liu, X. Zhang, Thermal Cracking of Hydrocarbon Aviation Fuels in Regenerative Cooling Microchannels, Energy Fuels. 27 (2013) 2563–2577.

B. Jin, K. Jing, J. Liu, X. Zhang, G. Liu, Pyrolysis and coking of endothermic hydrocarbon fuel in regenerative cooling channel under different pressures, J. Anal. Appl. Pyrol. 125 (2017) 117–126.

K.J. Johnson, B.J. Prazen, D.C. Young, R.E. Synovec, Quantification of naphthalenes in jet fuel with GC×GC/Tri-PLS and windowed rank minimization retention time alignment, J. Sep. Science. 27 (2004) 410–416.

K.J. Johnson, R.E. Synovec, Pattern recognition of jet fuels: comprehensive GC×GC with ANOVA-based feature selection and principal component analysis, Chem. Intel. Lab. Sys. 60 (2002) 225–237.

N.L. Johnson, A Simple Theoretical Approach to Cumulative Sum Control Charts, J. Amer. Stat. Assoc. 56 (1961) 835–840.

H.S. Johnston, C.E. Miller, B.Y. Oh, K.O. Patten, W.N. Sisk, Internal energy distributions from nitrogen dioxide fluorescence. 1. Cumulative sum method, J. Phys. Chem. 97 (1993) 9890–9903.

H. Kaiser, Quantitation in elemental analysis, Anal. Chem. 42 (1970) 24A–41a.

B. Kehimkar, J.C. Hoggard, L.C. Marney, M.C. Billingsley, C.G. Fraga, T.J. Bruno, R.E. Synovec, Correlation of rocket propulsion fuel properties with chemical composition using comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometry followed by partial least squares regression analysis, J. Chromatogr. A. 1327 (2014) 132–140.

B. Kehimkar, B.A. Parsons, J.C. Hoggard, M.C. Billingsley, T.J. Bruno, R.E. Synovec, Modeling RP-1 fuel advanced distillation data using comprehensive two-dimensional gas chromatography

coupled with time-of-flight mass spectrometry and partial least squares analysis, Anal Bioanal Chem. 407 (2015) 321–330.

W. Khummueng, J. Harynuk, P.J. Marriott, Modulation Ratio in Comprehensive Two-dimensional Gas Chromatography, Anal. Chem. 78 (2006) 4578–4587.

W. Khummueng, P.J. Marriott, The Nomenclature of Comprehensive Two-Dimensional Gas Chromatography: Defining the Modulation Ratio (MR), LCGC Europe. 22 (2009) 38–45.

H. Kimura, J. Goto, S. Yasuda, S. Sakurai, M. Yumura, D.N. Futaba, K. Hata, Unexpectedly High Yield Carbon Nanotube Synthesis from Low-Activity Carbon Feedstocks at High Concentrations, ACS Nano. 7 (2013) 3150–3157.

R.M. Kinghorn, P.J. Marriott, P.A. Dawes, Design and Implementation of Comprehensive Gas Chromatography with Cryogenic Modulation, J. High Resol. Chromatogr. 23 (2000) 245–252.

M.S. Klee, J. Cochran, M. Merrick, L.M. Blumberg, Evaluation of conditions of comprehensive two-dimensional gas chromatography that yield a near-theoretical maximum in peak capacity gain, J. Chromatogr. A. 1383 (2015) 151–159.

J.E. Knoll, Estimation of the Limit of Detection in Chromatography, J Chromatogr Sci. 23 (1985) 422–425.

F.D. Kopinke, G. Zimmermann, S. Nowak, On the mechanism of coke formation in steam cracking—conclusions from results obtained by tracer experiments, Carbon. 26 (1988) 117–124.

F.D. Kopinke, G. Zimmermann, G.C. Reyniers, G.F. Froment, Relative rates of coke formation from hydrocarbons in steam cracking of naphtha. 3. Aromatic hydrocarbons, Ind. Eng. Chem. Res. 32 (1993) 2620–2625.

P. Korytár, P.E.G. Leonards, J. de Boer, U.A.T. Brinkman, High-resolution separation of polychlorinated biphenyls by comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 958 (2002) 203–218.

J. Krupčík, P. Májek, R. Gorovenko, J. Blaško, R. Kubinec, P. Sandra, Considerations on the determination of the limit of detection and the limit of quantification in one-dimensional and comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1396 (2015) 117–130.

N. Kumar, A. Bansal, G.S. Sarma, R.K. Rawal, Chemometrics tools used in analytical chemistry: An overview, Talanta. 123 (2014) 186–199.

H.-H. Kuo, White Noise Distribution Theory, CRC Press, 1996.

H. Lander, A.C. Nixon, Endothermic fuels for hypersonic vehicles, Journal of Aircraft. 8 (1971) 200–207.

B.K. Lavine, J. Ritter, A.J. Moores, M. Wilson, A. Faruque, H.T. Mayfield, Source Identification of Underground Fuel Spills by Solid-Phase Microextraction/High-Resolution Gas Chromatography/Genetic Algorithms, Anal. Chem. 72 (2000) 423–431.

E.B. Ledford, C. Billesbach, Jet-Cooled Thermal Modulator for Comprehensive Multidimensional Gas Chromatography, J. High Resol. Chromatogr. 23 (2000) 202–204.

I.C. Lee, H.C. Ubanyionwu, Determination of sulfur contaminants in military jet fuels, Fuel. 87 (2008) 312–318.

A.C. Lewis, N. Carslaw, P.J. Marriott, R.M. Kinghorn, P. Morrison, A.L. Lee, K.D. Bartle, M.J. Pilling, A larger pool of ozone-forming carbon compounds in urban atmospheres, Nature. 405 (2000) 778–781.

G. Liu, Y. Han, L. Wang, X. Zhang, Z. Mi, Solid Deposits from Thermal Stressing of n-Dodecane and Chinese RP-3 Jet Fuel in the Presence of Several Initiators, Energy Fuels. 23 (2009) 356–365.

G. Liu, L. Wang, H. Qu, H. Shen, X. Zhang, S. Zhang, Z. Mi, Artificial neural network approaches on composition–property relationships of jet fuels based on GC–MS, Fuel. 86 (2007) 2551–2559.

Z. Liu, J. B. Phillips, Comprehensive Two-Dimensional Gas Chromatography using an On-Column Thermal Modulator Interface, J. Chromatogr. Sci. 29 (1991) 227–231.

Z. Liu, D.G. Patterson, M.L. Lee, Geometric Approach to Factor Analysis for the Estimation of Orthogonality and Practical Peak Capacity in Comprehensive Two-Dimensional Separations, Anal. Chem. 67 (1995) 3840–3845.

V.V. Lobodin, E.V. Maksimova, R.P. Rodgers, Gas Chromatography/Atmospheric Pressure Chemical Ionization Tandem Mass Spectrometry for Fingerprinting the Macondo Oil Spill, Anal. Chem. 88 (2016) 6914–6922.

G.L. Long, J.D. Wineforder, Limit of Detection: A Closer Look at the IUPAC Definition, Anal. Chem. 55 (1983) 712A–724A.

T.M. Lovestead, B.C. Windom, J.R. Riggs, C. Nickell, T.J. Bruno, Assessment of the Compositional Variability of RP-1 and RP-2 with the Advanced Distillation Curve Approach, Energy Fuels. 24 (2010) 5611–5623.

H. Lu, Q. Shi, J. Lu, G. Sheng, P. Peng, C.S. Hsu, Petroleum Sulfur Biomarkers Analyzed by Comprehensive Two-Dimensional Gas Chromatography Sulfur-Specific Detection and Mass Spectrometry, Energy Fuels. 27 (2013) 7245–7251.

J.M. Lucas, R.B. Crosier, Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM A Head Start, Technometrics. 24 (1982) 199–205.

J. Luong, X. Guan, S. Xu, R. Gras, R.A. Shellie, Thermal Independent Modulator for Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 88 (2016) 8428–8432.

F. Magagna, A. Guglielmetti, E. Liberto, S.E. Reichenbach, E. Allegrucci, G. Gobino, C. Bicchi, C. Cordero, Comprehensive Chemical Fingerprinting of High-Quality Cocoa at Early Stages of Processing: Effectiveness of Combined Untargeted and Targeted Approaches for Classification and Discrimination, J. Agric. Food Chem. 65 (2017) 6329–6341.

Markes International, Ltd. Application Note 537 (2016).

L.C. Marney, W. Christopher Siegler, B.A. Parsons, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-based Fisher-ratio software for improved feature selection analysis of comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry data, Talanta. 115 (2013) 887–895.

P. Marriott, R. Shellie, Principles and applications of comprehensive two-dimensional gas chromatography, Trends Anal. Chem. 21 (2002) 573–583.

P.J. Marriott, S.-T. Chin, B. Maikhunthod, H.-G. Schmarr, S. Bieri, Multidimensional gas chromatography, Trends Anal. Chem. 34 (2012) 1–21.

P.J. Marriott, R.M. Kinghorn, Longitudinally Modulated Cryogenic System. A Generally Applicable Approach to Solute Trapping and Mobilization in Gas Chromatography, Anal. Chem. 69 (1997) 2582–2588.

P.J. Marriott, P. Morrison, R.A. Shellie, M.S. Dunn, E. Sari, D. Ryan, Recent applications in multidimensional chromatography, LCGC Europe. 16 (2003) 23–31.

P.J. Marriott, R. Shellie, C. Cornwell, Gas chromatographic technologies for the analysis of essential oils, J. Chromatogr. A. 936 (2001) 1–22.

R.A.G. Marshall, Cumulative sum charts for monitoring of radioactivity background count rates, Anal. Chem. 49 (1977) 2193–2196.

J.J. McKetta, Encyclopedia of Chemical Processing and Design, CRC Press, 1990.

I.G. McWilliam, R.A. Dewar, Flame Ionization Detector for Gas Chromatography, Nature. 181 (1958) 760.

MIL-DTL-25576 E PROPELLANT ROCKET GRADE KEROSENE, (n.d.).

J.M. Miller, Chromatography, John Wiley & Sons, Inc., 2009.

S. Mitschke, W. Welthagen, R. Zimmermann, Comprehensive Gas Chromatography−Time-of-Flight Mass Spectrometry Using Soft and Selective Photoionization Techniques, Anal. Chem. 78 (2006) 6364–6375.

A.R. Mohan, S. Eser, Analysis of Carbonaceous Solid Deposits from Thermal Oxidative Stressing of Jet-A Fuel on Iron- and Nickel-Based Alloy Surfaces, Ind. Eng. Chem. Res. 49 (2010) 2722–2730.

R.E. Mohler, K.M. Dombek, J.C. Hoggard, E.T. Young, R.E. Synovec, Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry Analysis of Metabolites in Fermenting and Respiring Yeast Cells, Anal. Chem. 78 (2006) 2700–2709.

S. Mostafapour, H. Parastar, N-way partial least squares with variable importance in projection combined to GC × GC-TOFMS as a reliable tool for toxicity identification of fresh and weathered crude oils, Anal Bioanal Chem. 407 (2015) 285–295.

B. Munson, Chemical ionization mass spectrometry: ten years later, Anal. Chem. 49 (1977) 772A–775A.

M.S.B. Munson, F.H. Field, Chemical Ionization Mass Spectrometry. I. General Introduction, J. Am. Chem. Soc. 88 (1966) 2621–2630.

R.E. Murphy, M.R. Schure, J.P. Foley, Effect of Sampling Rate on Resolution in Comprehensive Two-Dimensional Liquid Chromatography, Anal. Chem. 70 (1998) 1585–1594.

A.M. Muscalu, M. Edwards, T. Górecki, E.J. Reiner, Evaluation of a single-stage consumable-free modulator for comprehensive two-dimensional gas chromatography: Analysis of polychlorinated biphenyls, organochlorine pesticides and chlorobenzenes, J. Chromatogr. A. 1391 (2015) 93–101.

J.S. Nadeau, R.B. Wilson, J.C. Hoggard, B.W. Wright, R.E. Synovec, Study of the interdependency of the data sampling ratio with retention time alignment and principal component analysis for gas chromatography, J. Chromatogr. A. 1218 (2011) 9091–9101.

M. Navarro-Reig, J. Jaumot, T.A. van Beek, G. Vivó-Truyols, R. Tauler, Chemometric analysis of comprehensive LC×LC-MS data: Resolution of triacylglycerol structural isomers in corn oil, Talanta. 160 (2016) 624–635.

A. Nijhuis, S. de Jong, B.G.M. Vandeginste, The application of multivariate quality control in gas chromatography, Chem. Intel. Lab. Sys. 47 (1999) 107–125.

A.B.J. Nix, S.J. Rowland, K.W. Kemp, D.W. Wilson, K. Griffiths, Internal quality control in clinical chemistry: A teaching review, Stat. Med. 6 (1987) 425–440.

Nomenclature, Symbols, Units and Their Usage in Spectrochemical Analysis-II. Data Interpretation, International Union of Pure and Applied Chemistry, Spectrochimica Acta Part B. 33 (1978) 241–245.

M.L. Oca, M.C. Ortiz, A. Herrero, L.A. Sarabia, Optimization of a GC/MS procedure that uses parallel factor analysis for the determination of bisphenols and their diglycidyl ethers after migration from polycarbonate tableware, Talanta. 106 (2013) 266–280.

B. Omais, M. Courtiade, N. Charon, D. Thiébaut, A. Quignard, M.-C. Hennion, Investigating comprehensive two-dimensional gas chromatography conditions to optimize the separation of oxygenated compounds in a direct coal liquefaction middle distillate, J. Chromatogr. A. 1218 (2011) 3233–3240.

J. Omar, M. Olivares, J.M. Amigo, N. Etxebarria, Resolution of co-eluting compounds of Cannabis Sativa in comprehensive two-dimensional gas chromatography/mass spectrometry detection with Multivariate Curve Resolution-Alternating Least Squares, Talanta. 121 (2014) 273–280.

L.S. Ott, A.B. Hadler, T.J. Bruno, Variability of The Rocket Propellants RP-1, RP-2, and TS-5: Application of a Composition- and Enthalpy-Explicit Distillation Curve Method, Ind. Eng. Chem. Res. 47 (2008) 9225–9233.

E.S. Page, Continuous Inspection Schemes, Biometrika. 41 (1954) 100–115.

E.S. Page, Cumulative Sum Charts, Technometrics. 3 (1961) 1–9.

H. Parastar, S. Mostafapour, G. Azimi, Quality assessment of gasoline using comprehensive two-dimensional gas chromatography combined with unfolded partial least squares: A reliable approach for the detection of gasoline adulteration, J. Sep. Science. 39 (2016) 367–374.

B.A. Parsons, L.C. Marney, W.C. Siegler, J.C. Hoggard, B.W. Wright, R.E. Synovec, Tile-Based Fisher Ratio Analysis of Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry (GC × GC–TOFMS) Data Using a Null Distribution Approach, Anal. Chem. 87 (2015) 3812–3819.

B.A. Parsons, D.K. Pinkerton, B.W. Wright, R.E. Synovec, Chemical characterization of the acid alteration of diesel fuel: Non-targeted analysis by two-dimensional gas chromatography coupled with time-of-flight mass spectrometry with tile-based Fisher ratio and combinatorial threshold determination, J. Chromatogr. A. 1440 (2016) 179–190.

P. de Peinder, T. Visser, R. Wagemans, J. Blomberg, H. Chaabani, F. Soulimani, B.M. Weckhuysen, Sulfur Speciation of Crude Oils by Partial Least Squares Regression Modeling of Their Infrared Spectra, Energy Fuels. 24 (2010) 557–562.

D.G. Peters, J.M. Hayes, G.M. Hieftje, Chemical separations and measurements: theory and practice of analytical chemistry, WB Saunders Company, 1974.

J.B. Phillips, R.B. Gaines, J. Blomberg, F.W.M. van der Wielen, J.-M. Dimandja, V. Green, J. Granger, D. Patterson, L. Racovalis, H.-J. de Geus, J. de Boer, P. Haglund, J. Lipsky, V. Sinha, E.B. Ledford, A Robust Thermal Modulator for Comprehensive Two-Dimensional Gas Chromatography, J. High Resol. Chromatogr. 22 (1999) 3–10.

J.B. Phillips, E.B. Ledford, Thermal modulation: A chemical instrumentation component of potential value in improving portability, Field Analyt. Chem. Technol. 1 (1996) 23–29.

J.B. Phillips, J. Xu, Comprehensive multi-dimensional gas chromatography, J. Chromatogr. A. 703 (1995) 327–334.

K.M. Pierce, J.C. Hoggard, J.L. Hope, P.M. Rainey, A.N. Hoofnagle, R.M. Jack, B.W. Wright, R.E. Synovec, Fisher Ratio Method Applied to Third-Order Separation Data To Identify Significant Chemical Components of Metabolite Extracts, Anal. Chem. 78 (2006) 5068–5075.

K.M. Pierce, B. Kehimkar, L.C. Marney, J.C. Hoggard, R.E. Synovec, Review of chemometric analysis techniques for comprehensive two dimensional separations data, J. Chromatogr. A. 1255 (2012) 3–11.

K.M. Pierce, S.P. Schale, Predicting percent composition of blends of biodiesel and conventional diesel using gas chromatography–mass spectrometry, comprehensive two-dimensional gas chromatography–mass spectrometry, and partial least squares analysis, Talanta. 83 (2011) 1254–1259.

D.K. Pinkerton, B.A. Parsons, T.J. Anderson, R.E. Synovec, Trilinearity deviation ratio: A new metric for chemometric analysis of comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry data, Anal. Chim. Acta. 871 (2015) 66–76.

D.K. Pinkerton, B.A. Parsons, R.E. Synovec, Method to determine the true modulation ratio for comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1476 (2016) 114–123.

M. Poliak, M. Kochman, A. Amirav, Pulsed flow modulation comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1186 (2008) 189–195.

J. Poloniecki, O. Valencia, P. Littlejohns, Cumulative risk adjusted mortality chart for detecting changes in death rate: observational study of heart surgery, BMJ. 316 (1998) 1697–1700.

T. Portolés, C. Sales, B. Gómara, J.V. Sancho, J. Beltrán, L. Herrero, M.J. González, F. Hernández, Novel Analytical Approach for Brominated Flame Retardants Based on the Use of Gas Chromatography-Atmospheric Pressure Chemical Ionization-Tandem Mass Spectrometry with Emphasis in Highly Brominated Congeners, Anal. Chem. 87 (2015) 9892–9899.

H. Potgieter, R. Bekker, J. Beigley, E. Rohwer, Analysis of oxidised heavy paraffininc products by high temperature comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 1509 (2017) 123–131.

B.J. Prazen, K.J. Johnson, A. Weber, R.E. Synovec, Two-Dimensional Gas Chromatography and Trilinear Partial Least Squares for the Quantitative Analysis of Aromatic and Naphthene Content in Naphtha, Anal. Chem. 73 (2001) 5677–5682.

S. Prebihalo, A. Brockman, J. Cochran, F.L. Dorman, Determination of emerging contaminants in wastewater utilizing comprehensive two-dimensional gas-chromatography coupled with time-of-flight mass spectrometry, J. Chromatogr. A. 1419 (2015) 109–115.

S.E. Prebihalo, K.L. Berrier, C.E. Freye, H.D. Bahaghighat, N.R. Moore, D.K. Pinkerton, R.E. Synovec, Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications, Anal. Chem. 90 (2018) 505–532.

M. Pursch, P. Eckerle, J. Biel, R. Streck, H. Cortes, K. Sun, B. Winniford, Comprehensive two-dimensional gas chromatography using liquid nitrogen modulation: set-up and applications, J. Chromatogr. A. 1019 (2003) 43–51.

B.D. Quimby, J.D. McCurry, W.M. Norman, Capillary flow technology for gas chromatography: Reinvigorating a mature analytical discipline, LCGC The Peak. (2007) 7–15.

J.R. Radović, K.V. Thomas, H. Parastar, S. Díez, R. Tauler, J.M. Bayona, Chemometrics-Assisted Effect-Directed Analysis of Crude and Refined Oil Using Comprehensive Two-Dimensional Gas Chromatography–Time-of-Flight Mass Spectrometry, Environ. Sci. Technol. 48 (2014) 3074–3083.

T. Rajalahti, O.M. Kvalheim, Multivariate data analysis in pharmaceutics: A tutorial review, Int. J. Pharm. 417 (2011) 280–290.

L. Ranzan, C. Ranzan, L.F. Trierweiler, J.O. Trierweiler, Classification of Diesel Fuel Using Two-Dimensional Fluorescence Spectroscopy, Energy Fuels. 31 (2017) 8942–8950.

B.C. Reaser, B.W. Wright, R.E. Synovec, Using Receiver Operating Characteristic Curves To Optimize Discovery-Based Software with Comprehensive Two-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry, Anal. Chem. 89 (2017) 3606–3612.

V.R. Reid, R.E. Synovec, High-speed gas chromatography: The importance of instrumentation optimization and the elimination of extra-column band broadening, Talanta. 76 (2008) 703–717.

G.C. Reyniers, G.F. Froment, F.-D. Kopinke, G. Zimmermann, Coke Formation in the Thermal Cracking of Hydrocarbons. 4. Modeling of Coke Formation in Naphtha Cracking, Ind. Eng. Chem. Res. 33 (1994) 2584–2590.

M.-F.S.G. Reyniers, G.F. Froment, Influence of Metal Surface and Sulfur Addition on Coke Deposition in the Thermal Cracking of Hydrocarbons, Ind. Eng. Chem. Res. 34 (1995) 773–785.
Richard Brereton, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Wiley, 2003.

Å. Rinnan, M. Andersson, C. Ridder, S.B. Engelsen, Recursive weighted partial least squares (rPLS): an efficient variable selection method using PLS, J. Chemometrics. 28 (2014) 439–447.

K. Robards, P.R. Haddad, P.E. Jackson, eds., Principles and Practice of Modern Chromatographic Methods, Academic Press, Boston, 2004.

E. Rubi, R.A. Lorenzo, M.C. Casais, R. Cela, Quality control in the routine analysis of methylmercury in biological and environmental materials using gas chromatography with electron capture detection, Appl. Organomet. Chem. 8 (1994) 677–686.

D. Ryan, R. Shellie, P. Tranchida, A. Casilli, L. Mondello, P. Marriott, Analysis of roasted coffee bean volatiles by using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, J. Chromatogr. A. 1054 (2004) 57–65.

P.J. Schoenmakers, J.L.M.M. Oomen, J. Blomberg, W. Genuit, G. van Velzen, Comparison of comprehensive two-dimensional gas chromatography and gas chromatography – mass spectrometry for the characterization of complex hydrocarbon mixtures, J. Chromatogr. A. 892 (2000) 29–46.

J.V. Seeley, Recent advances in flow-controlled multidimensional gas chromatography, J. Chromatogr. A. 1255 (2012) 24–37.

J.V. Seeley, F. Kramp, C.J. Hicks, Comprehensive Two-Dimensional Gas Chromatography via Differential Flow Modulation, Anal. Chem. 72 (2000) 4346–4352.

J.V. Seeley, F.J. Kramp, K.S. Sharpe, S.K. Seeley, Characterization of gaseous mixtures of organic compounds with dual-secondary column comprehensive two-dimensional gas chromatography (GC×2 GC), J. Sep. Science. 25 (2002) 53–59.

J.V. Seeley, N.J. Micyus, S.V. Bandurski, S.K. Seeley, J.D. McCurry, Microfluidic Deans Switch for Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 79 (2007) 1840–1847.

J.V. Seeley, N.E. Schimmel, S.K. Seeley, The multi-mode modulator: A versatile fluidic device for two-dimensional gas chromatography, J. Chromatogr. A. in press (2017).

J.V. Seeley, S.K. Seeley, Multidimensional Gas Chromatography: Fundamental Advances and New Applications, Anal. Chem. 85 (2013) 557–578.

J.V. Seeley, S.K. Seeley, Comprehensive two-dimensional gas chromatography with pattern modulation, J. Chromatogr. A. 1421 (2015) 114–122.

S. Shao, H. Zhang, Y. Wang, R. Xiao, L. Heng, D. Shen, Catalytic Pyrolysis of Biomass-Derived Compounds: Coking Kinetics and Formation Network, Energy Fuels. 29 (2015) 1751–1757.

R. Shellie, P. Marriott, P. Morrison, Concepts and Preliminary Observations on the Triple-Dimensional Analysis of Complex Volatile Samples by Using GC×GC−TOFMS, Anal. Chem. 73 (2001) 1336–1344.

R. Shellie, L. Mondello, P. Marriott, G. Dugo, Characterisation of lavender essential oils by using gas chromatography–mass spectrometry with correlation of linear retention indices and comparison with comprehensive two-dimensional gas chromatography, J. Chromatogr. A. 970 (2002) 225–234.

X. Shi, H. Li, Z. Song, X. Zhang, G. Liu, Quantitative composition-property relationship of aviation hydrocarbon fuel based on comprehensive two-dimensional gas chromatography with mass spectrometry and flame ionization detector, Fuel. 200 (2017) 395–406.

W.C. Siegler, J.A. Crank, D.W. Armstrong, R.E. Synovec, Increasing selectivity in comprehensive three-dimensional gas chromatography via an ionic liquid stationary phase column in one dimension, J. Chromatogr. A. 1217 (2010) 3144–3149.

W.C. Siegler, B.D. Fitz, J.C. Hoggard, R.E. Synovec, Experimental Study of the Quantitative Precision for Valve-Based Comprehensive Two-Dimensional Gas Chromatography, Anal. Chem. 83 (2011) 5190–5196.

A.E. Sinha, C.G. Fraga, B.J. Prazen, R.E. Synovec, Trilinear chemometric analysis of two-dimensional comprehensive gas chromatography–time-of-flight mass spectrometry data, J. Chromatogr. A. 1027 (2004) 269–277.

A.E. Sinha, K.J. Johnson, B.J. Prazen, S.V. Lucas, C.G. Fraga, R.E. Synovec, Comprehensive two-dimensional gas chromatography of volatile and semi-volatile components using a diaphragm valve-based instrument, J. Chromatogr. A. 983 (2003) 195–204.

A.E. Sinha, B.J. Prazen, C.G. Fraga, R.E. Synovec, Valve-based comprehensive two-dimensional gas chromatography with time-of-flight mass spectrometric detection: instrumentation and figures-of-merit, J. Chromatogr. A. 1019 (2003) 79–87.

H. Smith, R.D. Sacks, Column Selectivity Programming and Fast Temperature Programming for High-Speed GC Analysis of Purgeable Organic Compounds, Anal. Chem. 70 (1998) 4960–4966.

P.-H. Stefanuto, K.A. Perrault, L.M. Dubois, B. L'Homme, C. Allen, C. Loughnane, N. Ochiai, J.-F. Focant, Advanced method optimization for volatile aroma profiling of beer using two-dimensional gas chromatography time-of-flight mass spectrometry, J. Chromatogr. A. 1507 (2017) 45–52.

S.E. Stein, An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data, J. Am. Soc. Mass Spectrom. 10 (1999) 770–781.

B. Stiegemeier, M. Meyer, R. Taghavi, A Thermal Stability and Heat Transfer Investigation of Five Hydrocarbon Fuels, in: 38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, American Institute of Aeronautics and Astronautics, n.d.

X. Sun, C.M. Zimmermann, G.P. Jackson, C.E. Bunker, P.B. Harrington, Classification of jet fuels by fuzzy rule-building expert systems applied to three-way data by fast gas chromatography—fast scanning quadrupole ion trap mass spectrometry, Talanta. 83 (2011) 1260–1268.

R.E. Synovec, E.S. Yeung, Improvement of the limit of detection in chromatography by an integration method, Anal. Chem. 57 (1985) 2162–2167.

S. Tang, N. Shi, J. Wang, A. Tang, Comparison of the anti-coking performance of CVD TiN, TiO2 and TiC coatings for hydrocarbon fuel pyrolysis, Ceramics International. 43 (2017) 3818–3823.

P.Q. Tranchida, G. Purcaro, P. Dugo, L. Mondello, G. Purcaro, Modulators for comprehensive two-dimensional gas chromatography, Trends Anal. Chem. 30 (2011) 1437–1461.

J.L. Trenzado, J.S. Matos, L. Segade, E. Carballo, Densities, Viscosities, and Related Properties of Some (Methyl Ester + Alkane) Binary Mixtures in the Temperature Range from 283.15 to 313.15 K, J. Chem. Eng. Data. 46 (2001) 974–983.

R. Venkataraman, S. Eser, Characterization of Solid Deposits Formed from Short Durations of Jet Fuel Degradation: Carbonaceous Solids, Ind. Eng. Chem. Res. 47 (2008) 9337–9350.

H. van der Voet, Comparing the predictive accuracy of models using a simple randomization test, Chem. Intel. Lab. Sys. 25 (1994) 313–323.

R.E. Walpole, R.H. Myers, Probability and statistics for engineers and scientists, 2nd ed., MacMillan Publishing Co., Inc., New York, 1978.

J. Wang, M.-F. Reyniers, G.B. Marin, Influence of Dimethyl Disulfide on Coke Formation during Steam Cracking of Hydrocarbons, Ind. Eng. Chem. Res. 46 (2007) 4134–4148.

N.E. Watson, H.D. Bahaghighat, K. Cui, R.E. Synovec, Comprehensive Three-Dimensional Gas Chromatography with Time-of-Flight Mass Spectrometry, Anal. Chem. 89 (2017) 1793–1800.

N.E. Watson, B.A. Parsons, R.E. Synovec, Performance evaluation of tile-based Fisher Ratio analysis using a benchmark yeast metabolome dataset, J. Chromatogr. A. 1459 (2016) 101–111.

N.E. Watson, S.E. Prebihalo, R.E. Synovec, Targeted analyte deconvolution and identification by four-way parallel factor analysis using three-dimensional gas chromatography with mass spectrometry data, Anal. Chim. Acta. 983 (2017) 67–75.

N.E. Watson, W.C. Siegler, J.C. Hoggard, R.E. Synovec, Comprehensive Three-Dimensional Gas Chromatography with Parallel Factor Analysis, Anal. Chem. 79 (2007) 8270–8280.

R.L. Webster, P.M. Rawson, C. Kulsing, D.J. Evans, P.J. Marriott, Investigation of the Thermal Oxidation of Conventional and Alternate Aviation Fuels with Comprehensive Two-Dimensional Gas Chromatography Accurate Mass Quadrupole Time-of-Flight Mass Spectrometry, Energy Fuels. 31 (2017) 4886–4894.

B.A. Weggler, T. Gröger, R. Zimmermann, Advanced scripting for the automated profiling of two-dimensional gas chromatography-time-of-flight mass spectrometry data from combustion aerosol, J. Chromatogr. A. 1364 (2014) 241–248.

W. Welthagen, R.A. Shellie, J. Spranger, M. Ristow, R. Zimmermann, O. Fiehn, Comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry (GC × GC-TOF) for high resolution metabolomics: biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice, Metabolomics. 1 (2005) 65–73.

F. Westad, N.K. Afseth, R. Bro, Finding relevant spectral regions between spectroscopic techniques by use of cross model validation and partial least squares regression, Anal. Chim. Acta. 595 (2007) 323–327.

K.R. Williams, Colored Bands: History of Chromatography, J. Chem. Educ. 79 (2002) 922.

R.B. Wilson, B.D. Fitz, B.C. Mannion, T. Lai, R.K. Olund, J.C. Hoggard, R.E. Synovec, High-speed cryo-focusing injection for gas chromatography: Reduction of injection band broadening with concentration enrichment, Talanta. 97 (2012) 9–15.

R.B. Wilson, J.C. Hoggard, R.E. Synovec, High throughput analysis of atmospheric volatile organic compounds by thermal injection – isothermal gas chromatography – time-of-flight mass spectrometry, Talanta. 103 (2013) 95–102.

R.B. Wilson, W.C. Siegler, J.C. Hoggard, B.D. Fitz, J.S. Nadeau, R.E. Synovec, Achieving high peak capacity production for gas chromatography and comprehensive two-dimensional gas chromatography by minimizing off-column peak broadening, J. Chromatogr. A. 1218 (2011) 3130–3139.

B.C. Windom, T.J. Bruno, Assessment of the Composition and Distillation Properties of Thermally Stressed RP-1 and RP-2: Application to Fuel Regenerative Cooling, Energy Fuels. 25 (2011) 5200–5214.

J.D. Wineforder, Trace Analysis: Spectroscopic Methods for Elements, Wiley, New York, 1976.
S. Wold, K. Esbensen, P. Geladi, Principal component analysis, Chem. Intel. Lab. Sys. 2 (1987) 37–52.

S. Wold, J. Trygg, A. Berglund, H. Antti, Some recent developments in PLS modeling, Chem. Intel. Lab. Sys. 58 (2001) 131–150.

Z. Zeng, J. Li, H.M. Hugel, G. Xu, P.J. Marriott, Interpretation of comprehensive two-dimensional gas chromatography data using advanced chemometrics, Trends Anal. Chem. 53 (2014) 150–166.

Y. Zhang, S. Kajitani, M. Ashizawa, Y. Oki, Tar destruction and coke formation during rapid pyrolysis and gasification of biomass in a drop-tube furnace, Fuel. 89 (2010) 302–309.

K. Zhou, J. Jia, X. Li, X. Pang, C. Li, J. Zhou, G. Luo, F. Wei, Continuous vinyl chloride monomer production by acetylene hydrochlorination on Hg-free bismuth catalyst: From lab-scale catalyst characterization, catalytic evaluation to a pilot-scale trial by circulating regeneration in coupled fluidized beds, Fuel Process. Technol. 108 (2013) 12–18.

G. Zweig, J. Sherma, Paper Chromatography—Past, Present and Future, J Chromatogr Sci. 11 (1973) 279–283.

VITA

Christopher E. Freye was born in Huntsville Alabama in 1990 and grew up in Chattanooga, TN. After graduating from Notre Dame High School (Chattanooga, TN) in 2009, he attended the University of Tennessee at Knoxville where he graduate *cum laude* in 2013 with a degree in chemistry. Having completed his PhD, he will be taking a short vacation overseas before hopefully finding an interesting job.