The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:

Document Title:	**Statistical Assessment of the Probability of Correct Identification of Ignitable Liquids in Fire Debris Analysis**
Author(s):	**Michael Sigman, Ph.D., Mary Williams, M.S., Erin Waddell, Ph.D.**
Document No.:	**248564**
Date Received:	**January 2015**
Award Number:	**2009-DN-BX-K227**

This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this Federally-funded grant report available electronically.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Statistical Assessment of the Probability of Correct Identification of Ignitable Liquids in Fire Debris Analysis

Award No. 2009-DN-BX-K227

Author/ Principal Investigator

Michael E. Sigman, Ph.D.
Department of Chemistry and National Center for Forensic Science
University of Central Florida
PO Box 162367
Orlando, FL 32816-2367
407-823-6469

Co-Authors

Mary Williams, M.S.
National Center for Forensic Science
University of Central Florida
PO Box 162367
Orlando, FL 32816-2367

Erin Waddell, Ph.D.
National Center for Forensic Science
University of Central Florida
PO Box 162367
Orlando, FL 32816-2367

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Abstract

Identification of ignitable liquid residues in the presence of background interferences, especially those arising from pyrolysis processes, is a major challenge for the fire debris analyst. The proposed research will lead to a mathematical model that allows for the detection of an ignitable liquid in a fire debris sample and the classification of the ignitable liquid according to the ASTM E1618 classification scheme. The research will examine the influence of substrate pyrolysis and non-pyrolysis interferences on: (1) probability of correct prediction of the presence of an ignitable liquid in real and simulated fire debris samples (Type I and Type II error rates) and (2) probability of correct prediction of the associated ignitable liquid ASTM class and sub-class (heavy, medium or light) in positive samples. Potential alternative sub-groupings of ignitable liquids will be examined based on cluster analysis techniques. Models will be examined which are based on principal components analysis (PCA), linear discriminant analysis (LDA) and soft independent model classification analogy (SIMCA). The model will be developed from the summed ion spectra of nearly 500 ignitable liquid and 50 pyrolysis sample GC-MS data sets with ANOVA-assisted variable selection. Training data sets will be taken from the National Center for Forensic Science ignitable liquid and substrate pyrolysis databases. Simulated fire debris samples generated in the laboratory and samples from large-scale burns will also be employed in model testing. Model performance will be statistically evaluated by receiver operator characteristic analysis. The final model will be implemented in a software solution for forensic laboratory use.

This project proposed to investigate the development of a method for classifying fire debris GC-MS data sets as: (1) containing or not containing an ignitable liquid, (2) classifying any ignitable liquid that may be present under the ASTM E1618 classification scheme and (3)

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

estimating the statistical certainty of the answers to questions 1 and 2. The proposed approach is to build a mathematical model that can correctly classify GC-MS data from ignitable liquids and pyrolyzed substrates (wood, plastic, etc.). The model will then be applied to GC-MS data from laboratory-generated fire debris samples, as well as ignitable liquids and substrates that were not used to build the model. The classification success of the model will allow a determination of the statistical performance of the model by ROC analysis. The model will be developed based on the total ion spectrum, which has already shown a propensity for classifying a set of ignitable liquids drawn from multiple ASTM classes.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Contents

Executive Summary 1
The Problem 1
Purpose of the Research 1
Research Design 2
Findings and Conclusions 2
Implications for Policy and Practice 9
I. Introduction 10
Statement of the problem 10
Literature review 10
Current Methods of Fire Debris Analysis 11
Facilitating Inter-Laboratory Comparisons 12
Ignitable Liquid Groupings 15
Hierarchical Cluster 15
Self-Organizing Feature Maps (SOFM) 16
Automated Classification 18
Previous Work 19
Classification by LDA and QDA 24
Classification by SIMCA 25
Hypothesis or Rationale for the Research 28
II. Methods 29
Experimental Methods 29
Statistical Methods 31
Generation of the Data Sets 31
Total Ion Spectra Comparison 31
Cluster Analysis 32
Self-Organizing Feature Maps 33
Discriminant Analysis and SIMCA 34
Fire Debris Samples 35
Computational Details 38
Total Ion Spectra Comparisons 38
Cluster Analysis 39

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Self-Organizing Feature Maps (SOFM) 40
Discriminant Analysis Methods 41
Multi-Step Classification Procedure 42
Likelihood Ratio Test 45
Evaluation of Classifier Performance 45
SIMCA 46
Training 47
III. Results 48
Comparison of Total Ion Spectra 48
Grouping of Ignitable Liquids 54
Cluster Analysis 54
Self-Organizing Feature Maps 63
Discriminant Analysis 71
LDA and QDA 71
0% Substrate Contribution 73
20% Substrate Contribution 81
SIMCA 85
0% Substrate Contribution 88
20% Substrate Contribution 92
Results for Different Normalization Methods 95
Summed to One 95
Base Peak 97
Unit Vector 98
Training Results 101
IV. Conclusions 101
Discussion of Findings 101
Implications for Policy and Practice 103
Implications for Further Research 103
V. References 104
VI. Dissemination of Research Findings 109
Publications and Thesis 109
Patents 109
Presentations 109
Appendices 110

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Appendix I: Training Outline 110
Section 1: AMDIS 110
Section 2: R 111
Section 3: TIS 111
Section 4: Discriminant Analysis 111
Section 5: Putting it All Together 111
Final Exam 111
End of Course Survey 112
Appendix II: Demographic data for online and face-to-face applicants for the training 112
Appendix III: Face-to-Face Training Course Agenda 114

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Executive Summary

The Problem

Standardized practices for the extraction of ignitable liquid residues (1-5) and the analysis of the residues (6) have been established and are published by ASTM International. ASTM E1618-11 stipulates that analysis of extracts from fire debris samples to be analyzed by gas chromatography – mass spectrometry (GC-MS). Data analysis is comprised of visual pattern matching of the total ion chromatogram, extracted ion profiling, and target compound analysis. The ignitable liquid is classified into seven classes or a miscellaneous category as described within the standard (6). ASTM E1618-11 provides reliability of the scientific evidence consisting of testing, peer review, and general acceptance under the Daubert standard; however, it does not provide a means of determining error rates (7). The standard doesn't explicitly state how an analyst determines whether an ignitable liquid residue is present. It is assumed that if the analyst compares and matches chromatographic patterns of the unknown residue to a reference ignitable liquid that the residue is in fact an ignitable liquid. Therefore the analyst can state there was an ignitable liquid present in the fire debris. The analyst can't provide error rates using the method described because there is no mathematical basis and data analysis results are not conducive to statistical methods. In addition to the error rate problem, computational methods are often difficult to transition from academic research to operational forensic laboratories. As part of this research, the statistical methods developed in the academic research were transitioned to operational laboratories by online and face-to-face training.

Purpose of the Research

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The purpose of the research conducted under this award was to investigate potential ignitable liquid classification schemes as alternatives to ASTM E1618 and to develop a method for classifying fire debris GC-MS data sets as: (1) positive or negative for an ignitable liquid residue, (2) classifying any ignitable liquid residue that may be present under the ASTM E1618 or alternative classification scheme and (3) estimating the statistical certainty of the classification. Finally, the statistical methods developed under this grant were presented to fire debris analysts in operational laboratories through online and face-to-face training. The effectiveness of these two methods of transitioning research results to operational laboratories was evaluated.

Research Design

The research was conducted by utilizing data contained in the Ignitable Liquids Reference Collection (ILRC) database and the Substrate Database, which were developed as a collaboration between the National Center for Forensic Science (NCFS) at the University of Central Florida and the Scientific Working Group for Fire and Explosion (SWGFEX). The databases contain GC-MS data collected at the NCFS and reviewed by the SWGFEX Ignitable Liquids Database Committee to assign a consensus ASTM E1618 classification to each database record. The data analysis was accomplished using standard multivariate statistical techniques and calculations were performed using open-source statistical software that is freely available to forensic scientists. The results from the research have been published under peer-review and the best published classification methods were trained to a group of fire debris analysts in the United States and Europe by internet and face-to-face training.

Findings and Conclusions

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Fire debris is typically analyzed by gas chromatography–mass spectrometry (GC–MS) to test for the presence of ignitable liquid (IL) residues. Any residues extracted from fire debris are compared to reference ILs of known class following ASTM E1618-11 (6). Seven major classes of ILs are used: aromatic products (AR), gasoline (GAS), petroleum distillates (PD), isoparaffinic products (ISO), naphthenic-paraffinic products (NP), normal alkane products (NA), and oxygenated solvents (OXY). Samples are only assigned to a miscellaneous category (MISC) if they do not meet the criteria of one of the seven classes (6). The major classes of ILs, except for samples in the GAS class, are further divided based on the carbon range into light (L), medium (M), and heavy (H) subclasses. An IL or ignitable liquid residue (ILR) is determined to belong to one of the ASTM classes based on analyst interpretation of the total ion chromatogram (TIC), extracted ion profiles (EIP), and the presence of target compounds. Data analysis consists mainly of visual pattern recognition of the TIC with the interpretation criteria outlined in ASTM E1618-11 (6).

The examination of possible alternative classification schemes was based on analysis of the total ion spectrum (TIS), which is equivalent to the average mass spectrum across the chromatographic profile and has been demonstrated to identify and classify ignitable liquids with an accuracy comparable to the ability to identify organic compounds based on electron ionization mass spectra (8). Further characterization of the TIS constituted the first part of this work. The TIS for the 708 IL and SUB samples were binary encoded based on the four different transition levels (i.e., intensity relative to the base peak intensity of 1). The information content for transition levels 0.0001, 0.001, 0.01, and 0.1 were calculated, to be: 114.2, 125.1, 121.9, and 56.53 bits, respectively. These results are consistent with those found by Grotch for approximately 3,000 organic compounds (9). The number of disagreeing ions between each of

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

the 250,278 unique pairwise comparisons of the 708 IL and SUB samples was determined. The average observed number of disagreeing ions for transition levels 0.0001, 0.001, 0.01, and 0.1 are: 52.61, 58.16, 56.08, and 23.69 ions, respectively. Out of the 250,278 unique pairwise comparisons of the 708 IL and SUB samples at the 0.01 transition level, only 20 comparisons resulted in zero ion disagreements (approximately 8 x 10-3 %, which extrapolates to approximately 80 indistinguishable spectra in 106 comparisons). Grotch found six indistinguishable spectra in 106 comparisons for pure organic compounds (2). Each of the 20 indistinguishable comparisons in this work occurred between IL samples with the same ASTM class (or subclass). These results demonstrate that even when binary encoded at the transition levels studied in this work, the "mass spectrum is a very specific chemical signature" (9), and while 80 indistinguishable pairs of spectra for each 106 comparisons is significantly larger than three indistinguishable pairs, the binary encoded TIS comprise a significantly unique signature for complex mixtures, such as ignitable liquids.

The ASTM classifications of IL defined in E1618 are based primarily on product class and petroleum refining methods. It is worth considering other possible groupings or classification schemes for these liquids that may be dictated by the chemistry and/or mass spectral behavior, without prior bias based on ASTM E1618. Once groupings have been obtained, however, it is interesting to compare these grouping to those defined in ASTM E1618 in order to consider any new information or further understanding of the emerging classifications. Investigation of groupings based on the TIS of ignitable liquids and substrate pyrolysis samples was undertaken by two methods, hierarchical clustering and self-organizing feature maps. The TIS for 30 – 200 m/z were compared for each of 445 IL and 88 SUB samples and clustered by hierarchical cluster analysis with optimal leaf ordering. The correlation distance, average linkage methods were used

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

in the comparisons. The resulting clusters were found to strongly overlap with the ASTM E1618 defined classes, giving significant support to currently followed ignitable liquid classifications protocols. The IL corresponding to well-defined classes (i.e., excluding the substrates, miscellaneous and oxygenates) clustered into two distinct groups, one consisting of the aromatic and gasoline products and the other consisting of the aliphatic hydrocarbon containing classes. Within each of these two major clusters, sub-clustering was observed along the lines of the ASTM E1618 class definitions.

The self-organizing feature map made use of 30 ion intensities from the TIS to group similar ignitable liquids into neighborhoods (i.e. clusters of liquids sharing common features). The cluster structure of the high-dimensional input data may be visualized in a two-dimensional unified distance matrix (or U-matrix) which shows the normalized distances between the weight vector of each neuron and neighboring neurons Neighborhoods of low similarity in close proximity show a distinct line of demarcation when viewed on the U-matrix. The general grouping of neurons can be viewed in the context of ASTM classes by projecting class membership onto the U-matrix to determine if IL from the same ASTM class group together. The U-matrix for the calculated SOFM contained cluster borders which separate the cluster of neurons associated with AR, GAS, and OXY ASTM classes from those of ISO, NA, NP, and PD. Separation was also observed for the neurons associated with ISO and NA classes from those of PD and NP. As with the hierarchical clustering, the self-organizing feature map results reinforced the use of ASTM E1618 classes to define commercially available ignitable liquids. The intensity of the 30 ions used in the self-organizing calculation projected onto the U-matrix reveal the neighborhoods where they made major contributions and the results also reinforced the underlying scientific basis for ASTM E1618. Thirty three samples designated as

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

miscellaneous (MISC), which have spectral characteristics of multiple ASTM classes, were projected onto the U-matrix. Of these samples, 22 projected to neurons associated with one of the ASTM classes comprising the liquid or to empty neurons adjacent to a neighborhood associated with one of the classes comprising the liquid. Seven samples designated as OXY, which contain mixtures of oxygenated compounds with ASTM classs (i.e. LPD+acetone), were projected onto the U-matrix. Projection resulted in four of the samples being assigned to a neuron associated with the associated ASTM class, while one sample (LPD+acetone) is assigned to an empty neuron adjacent to both OXY and LPD associated neurons. These results reinforce not only the ASTM classification system, but the use of the TIS and extracted ion spectra to represent ignitable liquid classes as defined by ASTM E1618. The TIS were further utilized in the classification of ignitable liquids and substrate pyrolysis samples using multivariate statistical methods.

Linear and quadratic discriminant analysis methods were applied to the TIS of IL and substrate pyrolysis (SUB) samples to determine if these statistical methods could reliably classify samples as positive or negative for ILR. Samples that were determined to be positive for ILR were further classified through a stepwise process of binary classifications into the corresponding ASTM classes, excluding the oxygenated class and the miscellaneous category. In each case, the covariance matrices of the two classes were found to differ, making the quadratic discrimination technique more appropriate. Classification models were generated using a training data set that contained only IL and SUB samples, and with a data set containing IL, SUB, IL+20%SUB and mixtures of SUB samples. The true positive, false positive, true negative, false negative, precision and accuracy rates were calculated for the classification as positive or negative for ILR based on cross validation where 20% of each class was withheld randomly without replacement

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

for each of 100 iterations. The classification methods were also tested on fire debris data where the "ground truth" was known based on either proximity of the sampling location to the pour trail or by data inspection by an analyst with full knowledge of the IL used in the burn. The QDA model produced the lowest false positive rate (6%) in cross-validation for the model containing 20% substrate contribution. Similarly, QDA gave the lowest false positive rate (8.9%) for the fire debris samples with the class assigned based on proximity to the pour; however, the false positive rate increased when the fire debris sample designations were based on the analyst. The cross-validation results demonstrate that the QDA model can maintain low false positive rate for data sets that resemble the data used to create the model. The fire debris results, with designations based on the pour, demonstrate a lower false positive rate than when the designations were based on the analyst. The precision of the methods ranged from 85 – 97% and the accuracy was in the range of 82 – 92%. Similar classification was performed using soft independent model classification analogy (SIMCA) on the same data sets. SIMCA based on the 20% substrate contribution data set gave very good results, with a 94.2% true positive rate and 5.1% false positive rate with greater than 94% precision and accuracy. SIMCA detection of the presence of IL trace in fire debris samples was in good agreement with the analyst, showing greater than 95% true positive; however, the false positive rate for fire debris increased to approximately 15%.

The classification results are somewhat dependent on the method used to normalize the TIS for model development and fire debris classification. The methods investigated included adjusting all of the intensities in each TIS so they sum to one, setting the maximum intensity in each TIS equal to one ("base peak"), and normalizing each TIS to a vector length of one. Classification by the SIMCA method and normalization by the "base peak" method resulted in the highest true

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

positive rates and lowest false positive rates for the cross-validation test set. These methods did not perform as well for the fire debris samples and reflected a false positive rate that was nearly double the rate obtained using the "summed to one" method. When considering the QDA classification method, low false positive rates were obtained for the model and fire debris samples when using the "base peak" method. The best overall performance was observed for QDA classification using the "base peak" method. The normalization method giving the highest correct classification rate for each statistical method should be utilized.

In subsequent steps of the classification into ASTM classes for those samples that were classified as positive for ILR, the LDA and QDA performance was varied depending on the data normalization method applied prior to model development. Nonetheless the results for most classifications were greater than 80%, with some problems observed for discrimination between highly similar ASTM classes (i.e., PD vs. NP and NA vs. ISO). The results of these classifications are considered less important forensically than the classification of a sample as positive or negative for ILR. The correct classification rates for samples that were positive or negative for ILR are encouraging and reflect the possibilities for introducing statistical methods into the analysis of fire debris in order to bring this area of forensic analysis into closer alignment with the Daubert requirement.

Irrespective of the quality of the research results from this grant or others, forensic science will remain unchanged if the research is not transitioned from academic laboratories to operational laboratories. As a final part of this research program, the most successful discrimination method (QDA with base peak normalization) was presented to operational laboratory fire debris analysts by both online and face-to-face training motifs. The online and face-to-face training required a maximum of eight hours to complete. An international group of 35 fire debris analysts

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

volunteered for the course and only 25 (71%) participants (9 U.S. and 16 European) finished the course online. A second group of 15 U.S. fire debris analysts enrolled to participate in the face-to-face training. A total of 14 (93%) of the 15 participants were actually trained face-to-face. This difference in completion rates (71% versus 93%) is not statistically significant (α=0.13) at a power of 0.8, which is often considered satisfactory for a small sample size. The average score on the final examination in the course was 92% for the online participants and 90% for the face-to-face participants. This difference is not statistically significant and an indication that the methods can be trained either online or face-to-face. The demographic data for both sets of participants (online and face-to-face were fairly comparable (see Appendix II at the end of this report), with the most significant differences being that the face-to-face participants were older, self-reported to be more proficient in English and had more years of experience in fire debris analysis.

Implications for Policy and Practice

The results of this research provide the first large-scale demonstration of statistically reliable classification rates for fire debris as positive or negative for ignitable liquid residue. Fire debris analysis methods with known error rates meet the Daubert requirements and may someday be required under existing rules of evidence. The research results will help to drive policy change to improve the practice of forensic fire debris analysis.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

I. Introduction

Statement of the problem

Standardized practices for the extraction of ignitable liquid residues (1-5) and the analysis of the residues (6) have been established and are published by ASTM International. ASTM E1618-11 stipulates that analysis of extracts from fire debris samples should be analyzed by gas chromatography – mass spectrometry (GC-MS). Data analysis is comprised of visual pattern matching of the total ion chromatogram, extracted ion profiling, and target compound analysis. The ignitable liquid is classified into seven classes or a miscellaneous category as described within the standard (6). ASTM E1618-11 provides reliability of the scientific evidence consisting of testing, peer review, and general acceptance under the Daubert standard; however, it does not provide a means of determining error rates (7). The standard doesn't explicitly state how an analyst determines whether an ignitable liquid residue is present. It is assumed that if the analyst compares and matches chromatographic patterns of the unknown residue to a reference ignitable liquid that the residue is in fact an ignitable liquid. Therefore the analyst can state there was an ignitable liquid present in the fire debris. The analyst can't provide error rates using the method described because there is no mathematical basis and data analysis results are not conducive to statistical methods. The research conducted under this award investigated the development of a method for classifying fire debris GC-MS data sets as: (1) positive or negative for an ignitable liquid residue, (2) classifying any ignitable liquid residue that may be present under the ASTM E1618 classification scheme and (3) estimating the statistical certainty of the answers to questions 1 and 2.

Literature review

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Current Methods of Fire Debris Analysis

Fire debris is typically analyzed by gas chromatography–mass spectrometry (GC–MS) to test for the presence of ignitable liquid (IL) residues. Any residues extracted from fire debris are compared to reference ILs of known class following ASTM E1618-11 (6). Seven major classes of ILs are used: aromatic products (AR), gasoline (GAS), petroleum distillates (PD), isoparaffinic products (ISO), naphthenic-paraffinic products (NP), normal alkane products (NA), and oxygenated solvents (OXY). Samples are only assigned to a miscellaneous category (MISC) if they do not meet the criteria of one of the seven classes (6). The major classes of ILs, except for samples in the GAS class, are further divided based on the carbon range into light (L), medium (M), and heavy (H) subclasses. An IL or ignitable liquid residue (ILR) is determined to belong to one of the ASTM classes based on analyst interpretation of the total ion chromatogram (TIC), extracted ion profiles (EIP), and the presence of target compounds. Data analysis consists mainly of visual pattern recognition of the TIC with the interpretation criteria outlined in ASTM E1618-11 (6). Data interpretation is subject to the skill and experience of the analyst (10); however, an automated systematic approach to the analysis of fire debris samples was reported by Bertsch as early as 1988, one year preceding the formation of the ASTM Committee E 30 on Forensic Science (11). The method proposed by Bertsch was similar to the procedures subsequently adopted in ASTM E1618, making use of extracted ion profiles. Notably, Bertsch's paper was out of step with common practices of the time by promoting a computer automated approach that "requires only minimal interaction with the analyst." ASTM Method E1618 advocates the use of libraries and target compounds to assist in the pattern recognition of different classes of ignitable liquids (6). Detection of ignitable liquids in fire debris on the basis of target compounds was demonstrated by Keto in 1991, prior to the adoption of the first version

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

of E1618 in 1994 (12). Lennard later demonstrated a GC/MS database for ignitable liquids based on the use of target compounds (13). In 1997, the National Center for Forensic Science (NCFS) and the Technical Working Group for Fire and Explosion (TWGFEX) began a collaboration that established the Ignitable Liquid Reference Collection and Database (ILRC), which has grown to nearly 1,000 records that are freely available online for fire debris analysts (14).

Facilitating Inter-Laboratory Comparisons

The GC-MS is a useful analytical instrument for the analysis of complex mixtures and the resulting data are used for the identification of compounds based on their retention times in the total ion chromatograms (TICs) and mass spectra. However, analyses by laboratories with different instruments or varying methods generally result in retention time shifts. Peak alignment methods or conversion using retention indices may be performed; however, this is time consuming and does not readily lend itself to automation (8). A simpler method for comparing complex samples analyzed with different GC-MS methods involves analysis of the total ion spectrum (TIS). The TIS is defined as the average mass spectrum across the chromatographic profile (8); therefore, it contains the mass spectral information for the data set, independent of the chromatographic time component.

The question of whether or not TIS with a specified mass range contains enough information for identification of complex mixtures, such as ignitable liquids, has previously been addressed (8). The question of uniqueness of electron ionization mass spectra was also addressed for 3,000 individual organic molecules in 1970 by Grotch using information theory (9). In Grotch's work, the mass spectral data were binary encoded so a mass spectral peak above a specific transition level relative to the base peak (intensity of 1) was encoded as a "1" and

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

peaks with intensities below the transition level were encoded as a "0" (9). He pointed out that analysis of binary encoded mass spectral data with a mass range of 200 amu could theoretically encode 2^{200} different compounds; however, this value would decrease as the correlation between the ions present in the spectra increased (9). This behavior would be observed with isotopes or ions commonly observed for different types of compounds (*i.e. m/z* 43, 57, 71, etc. for *n*-alkane containing compounds). Grotch found that even binary encoded mass spectral data with mass range *m/z* 12-200 contained "highly specific signatures" (9). This principle was tested on the analysis of 440 commercially available IL samples by Sigman *et al.* in 2008 (8). When the mass spectral data, with a mass range of *m/z* 30-350, were binary encoded based on a one percent transition, an average of 50 differing ions were observed for the 96,580 unique pairwise comparisons. The 440 IL samples in the study were obtained from the ILRC database (2, 8). Further characterization of the TIS for the description of ignitable liquids and substrate pyrolysis products was examined in the research conducted under this award.

The following is a brief computational background regarding the binary encoding of spectra and the calculation of information content used by Grotch, and applied to TIS in our studies. Each mass spectrum or TIS is binary encoded by first setting a transition level for peak detection. If the intensity of an ion in the TIS is greater than or equal to the transition level, the peak is assigned an intensity of 1; otherwise, it is assigned a value of 0. For each transition level, the total number of differing bits was calculated for each pairwise sample comparison. Formally, each ion in the pair of the binary encoded TIS is compared as defined by the function $F(a_i,b_i)$ in Equations 1 and 2, and the observed number of disagreeing ions ($D_{L(obs)}$) for a two sample comparisons may be calculated by Equation 3 (9), where L is defined as the number ions in the spectrum.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

$$F(a_i, b_i) = 0 \quad \text{if } a_i = b_i \tag{1}$$

$$F(a_i, b_i) = 1 \quad \text{if } a_i \neq b_i \tag{2}$$

$$D_{L\,(obs)} = \sum_{i=1}^{i=L} F(a_i, b_i) \tag{3}$$

The average observed number of disagreeing ions ($\overline{D}_{L\,(obs)}$) was calculated for all of the unique pairwise comparisons of the samples at each of the transition levels. The average observed number of disagreeing ions was determined according to Equation 4:

$$\overline{D}_{L\,(obs)} = \frac{D_L}{P} \tag{4}$$

where *P* is defined as the number of pairwise comparisons. The number of pairwise comparisons for samples with the same designation was found (9) according to Equation 5:

$$P = \frac{N(N-1)}{2} \tag{5}$$

where *N* are the number of samples (9). In our case, when samples belong to the same class, P is calculated as in Equation 5; however, the numbers of pairwise comparisons for those samples that have differing designations are calculated according to Equation 6:

$$P = N_1 N_2 \tag{6}$$

where N_1 is the number of samples with one designation and N_2 is the number of samples with the second designation.

The information content for *L* ions (H_L), of the data may be expressed quantitatively by the information "entropy" that is measured in "bits"(9) and may be determined using Equation 7

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

(9), where M refers to the number of levels, which is two in the case of binary encoded data, and p_{ik} corresponds to the probability of a peak intensity for ion i falling in level k.

$$H_L = -\sum_{i=1}^{L}\sum_{k=1}^{M} p_{ik} \log_2 p_{ik} \tag{7}$$

The information content is maximized and equal to L when the chosen transition level is such that it is equally likely for a peak in the TIS to be present or not (9).

Ignitable Liquid Groupings

Hierarchical Cluster

The TIS allows for an approach to the clustering of ignitable liquids that is outside or independent of the framework of ASTM E1618 (6). Hierarchical cluster analysis of lighter fluids and medium petroleum distillates has previously been performed using the TIC; however, retention times of the inter-laboratory TIC will vary (8), as described above. Agglomerative hierarchical cluster analysis has previously been used to obtain rational groupings based on mass spectral data (15-18). Hierarchical cluster analysis is an unsupervised learning method which means that no prior knowledge of group membership is required (19). The agglomerative process begins by considering each object individually, then sequentially grouping the most similar objects, and lastly combining similar groups to form new groups. Ideally, objects within a group should be more similar to one another than to objects in other groups (20). The results are typically shown as a dendrogram. The dendrogram demonstrates how the objects or groups are connected using lines with lengths that reflect the distances between the objects or groups.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Agglomerative hierarchical cluster analysis was used in the research conducted under this award to reassess the groupings of ignitable liquids based on the chemistry of the groups, as reflected in the TIS of the liquids.

Self-Organizing Feature Maps (SOFM)

Visual pattern recognition, as required under ASTM E1618, is dependent on the interpretation of the chromatographic data by the analyst (8, 21) and does not readily lend itself to automation. Non-subjective chemometric techniques, which involve the extraction of information from chemical data, have been explored as a means of obviating the subjective nature of visual pattern recognition (22). Many chemometric techniques allow for data visualization by reducing the dimensionality of the data (23) and clustering or grouping of data based on similar characteristics (24). Kohonen's SOFMs also known as self-organizing maps (SOMs), is an artificial neural networks technique designed as a method for abstraction, clustering, and visualization of high-dimensional data (25-27). This is usually accomplished through nonlinear mapping onto a two-dimensional grid space of a predefined number of neurons, where each neuron contains a weight vector that is comprised of the same number of components as the number of variables (or dimensionality) of the input space (28). The weight vectors are adjusted through an unsupervised learning process resulting in the neurons being arranged according to patterns in the input signal (26). When a sample is introduced to the map, a distance is calculated between the sample's variable vector and the weight vector for each neuron, where the Euclidean distance is often used. The neuron with the smallest distance to the sample is determined to be the "winning neuron", and the weight vector for that neuron is adjusted to be more similar to the sample vector. The weight vectors of neighboring neurons that

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

are a given distance from the winning neuron are also updated, which preserves the neighborhood relationships of the data within the input space (28). This training process is repeated for each of the samples, resulting in the completion of one full cycle or epoch, and this process is repeated iteratively for a predefined number of epochs (29). A batch algorithm may also be used (25, 30). The amount of adjustment to the weight vectors and the neighborhood size are varied throughout the training process (31). In the last step of training, the input data is re-introduced to the grid, and the samples are mapped onto the winning neurons (29). This process results in a two-dimensional map of neurons, where sample data is clustered based on similarity while preserving the neighborhood relationships of the data (25, 30, 32).

SOFMs have previously been used for clustering and/or classifying samples analyzed by GC-MS. Strawberry varieties were studied using solid-phase micro-extraction GC-MS data and clustered according to type by SOFM (33). Crude oil (34) and weathered crude oil samples (29) were geographically classified. Weathered and unweathered lighter fluids were classified based on manufacturer (21), and classification according to product type and brands were performed for weathered and unweathered medium petroleum distillates (35). A set of 150 ions resulting from pyrolysis mass spectrometry was analyzed with SOFMs to classify plant seeds (31). While SOFMs have been used for classification, Kohonen points out that self-organizing feature maps are a beneficial unsupervised method for clustering, visualization, and abstraction, but are not meant for statistical pattern recognition (25). A method that is "particularly suitable for statistical pattern recognition" is a supervised version of SOFM known as Learning Vector Quantization (LVQ) (25). In addition to clustering, a SOFM allows visual associations between individual variables and clustered samples in order to determine feature attributes for natural clusters within the input data space.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

In the research conducted under this award, 313 ignitable liquids from varying ASTM classes were grouped using the unsupervised SOFM technique, and ignitable liquid residues were analyzed based on the SOFM model. The natural clusters of the input data and the relationships between these clusters and their spectral variables were examined with respect to the ASTM class designations of the ILs. In this portion of the research, extracted ion spectra (EIS) comprised of a limited set of ions that were identified from Table 2 of ASTM E1618 were utilized (6). The EIS of select samples designated as MISC or OXY, as well as ignitable liquid residues from fire debris samples, were projected onto the SOFM demonstrating the similarities and differences between the variables of the newly projected data compared to those of the data used to train the SOFM

Automated Classification

As described above, current practices in fire debris analysis involve using the total ion chromatogram (TIC), extracted ion profiles (EIP), and target compound analysis (TCA) to identify the presence of an ignitable liquid and assign it to the correct ASTM class (10, 36, 37). The methodology outlined by ASTM E1618 relies on visual pattern recognition to identify the presence of an ignitable liquid residue in a fire debris sample. While this method can work well for neat samples of ignitable liquids, classification of post-burn samples may be complicated due to the presence of substrate pyrolysis products and the evaporation of more volatile components (weathering) (38). The volume of ignitable liquid(s) used in a fire, the placement of the liquid(s) (i.e., on wood, carpet, etc.), and the temperature reached during the fire are among the variables that can influence the ratio of ignitable liquid residue-to-substrate pyrolysis products in a sample and the analyst's interpretation of the results (39). These factors can alter the chromatographic

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

and spectral data of the fire debris samples, thus complicating the classification of ignitable liquid residue (38, 40). Ultimately, it is impossible (and unnecessary) to quantitatively determine the ratio of ignitable liquid residue-to-substrate pyrolysis products, and a positive determination of ignitable liquid residue in fire debris samples requires a strong or clear chromatographic pattern that is representative of an ignitable liquid from a given ASTM class.

Visual pattern recognition can prove to be time consuming and possibly subjective for large sample sets. Applying multivariate statistical methods to fire debris samples can potentially improve the analyst's ability to identify the presence of ignitable liquid residue, discriminate between similar samples, and reduce the analysis time of the sample. In the research conducted under this award, multivariate classification methods were investigated for large data sets, and the results include statistically defensible error rates for classification of a fire debris sample as positive or negative for ignitable liquid residue.

Previous Work

Peer-reviewed literature that has attempted to address the questions of ignitable liquid detection and classification has focused primarily on the problem of classification, in many cases under the assumption or knowledge that an ignitable liquid was present in the sample. A limited number of reports have addressed the influence of interference from "pyrolysis products" and other sources on the classification of ignitable liquids. Pattern recognition and classification through the use principal components analysis (PCA), canonical variate analysis (CVA) and artificial neural networks (ANN) have been addressed. PCA is an unsupervised (i.e. *a priori* classification or groupings <u>is</u> <u>not</u> required) dimension reduction methodology. Application of PCA may show clustering of data but clustering is not guaranteed, whereas CVA is designed to

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

give optimal clustering by maximizing between-group variance and minimizing within-group variance. CVA is a supervised (i.e. *a priori* classification or groupings is required) data analysis method. A related method is soft independent modeling of class analogy (SIMCA) wherein PCA is performed on data sets for samples constituting different classes and the models applied to unknown samples with the use of a distance metric to assess class membership. The ANN is a very different approach to pattern recognition. There are many implementation strategies for ANNs, some of which utilize supervised learning and some utilize unsupervised learning. The paradigm most commonly applied to fire debris analysis is the back-propagation ANN. This ANN utilizes supervised training on a series of adjustable weights connecting the input, hidden, and output layers. The number of training sets is recommended to exceed the number of adjustable weights by a factor of three, otherwise the ANN will "memorize" the input data and fail to "extrapolate" when presented with new data (41). Since the number of adjustable weights is typically large, the large size of the training data set is a significant impediment to the implementation of an ANN solution. These previous works are described in following paragraphs.

Tan et al. has investigated the use of PCA and SIMCA for the classification of accelerants (36). The research examined 51 ignitable liquids from five classes (as designated by ASTM E1618-94 from 1994). Some ignitable liquids were soaked into substrates (wood, carpet, etc.) and recovered, while some substrates were evaporated to test the effects of weathering. Polyolefin carpet samples that had been charred in open air were determined to not produce any products that interfered with the data analysis; however, pyrolysis of the carpet led to multiple products that can obscure and accelerant pattern and make accelerant identification difficult. Data sets were analyzed by summing the total ion intensity in 19 time increments across the

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

chromatographic profile. SIMCA was determined to result in perfect classification of test samples. Many other groups have approached this problem, while limiting themselves to narrower range of ignitable liquid classes.

Sandercock has utilized PCA with subsequent linear discriminant analysis (LDA) of the PCA scores to discriminate between samples of unevaporated and evaporated gasoline based on the C_0- to C_2-alkylnaphthalenes in each sample(42, 43). In the second study, 35 gasoline samples were shown to form 18 groups based on the C_0- to C_2-alkylnaphthalenes, irrespective of the extent of evaporation. A related study used the peak areas from 44 target compounds in an attempt to classify gasoline samples as either premium or regular and to sub-classify the gasoline samples into winter or summer sub-groups (10). This study resulted in 80-93% correct classification of the gasoline samples as premium or regular, but the correct classification dropped to 48-62% when the gasoline samples were further sub-classified into summer/winter formulations. A back-propagation ANN was reported to give 97% correct classification down to the summer/winter sub-classification. The best performing network was found to have 44 input nodes, 18 hidden nodes and 2 output nodes. The ANN model had 1,580 adjustable weights and ideally would have required 4,740 unique training data sets. The strong performance by the ANN in this case may reflect memorization of the data set. In another multivariate approach, Petraco has examined gasoline samples from casework using PCA, CVA and orthogonal canonical variate analysis (OCVA) (44). The GC-MS data from liquid gasoline samples was analyzed using the peak intensities of 15 target compounds. CVA and OCVA of the data gave good classification of the 20 samples analyzed and allowed their discrimination given the *a priori* knowledge of their groups.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Synovec has taken an interesting approach to the analysis of GCxGC/MS data of jet fuel samples (45). Samples were assigned to groups based on fuel type (JP-7, JP-8, etc.) and the resulting GCxGC data were analyzed by ANOVA to determine those parameters (retention times on each GC column) which gave large *f* values (ratio of inter- to intra-class variance). The selected parameters were used to perform PCA, the results of which showed excellent grouping of the samples by fuel type. PCA of mixed samples were shown to give scores falling along the lines connecting the scores of the pure samples used to prepare the mixed samples. In a related approach, Borusiewicz has analyzed kerosene and diesel samples by first using MANOVA with a Tukey post-hoc test to select the chromatographic peaks that would be used in a PCA attempt at clustering (38). A set of only four variables (linear hydrocarbons C11, C15, C17 and C23) were chosen for PCA. The probability of association was assessed by a likelihood ratio and the pairwise comparison of 12 samples resulted in three Type II errors from 66 unique comparisons while holding the Type I error at 5%. Likelihood ratios have been examined elsewhere for the evaluation of trace evidence questioned – known comparisons (46). Hupp was also able to observe clustering between diesel fuels from different manufacturers using PCA analysis of chromatographic data and without prior statistical analysis of the variables (37).

Previous research from Sigman et al. has demonstrated the use of covariance mapping as a method for identification of the gas chromatography-mass spectrometry (GC-MS) data from complex mixtures often encountered in fire debris analysis, and a method of rapidly performing a numeric search of an ignitable liquid library to find a nearest match (47). One advantage of the covariance method is that it overcomes many of the problems associated with lab-to-lab variations in analyte chromatographic retention time. Covariance mapping has also been demonstrated to provide a method for discriminating between fresh gasoline samples at a known

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

level of statistical significance. While the covariance mapping approach fully utilizes the information-rich data sets available based on current GC-MS analysis methods, map calculation is somewhat CPU and memory intensive.

Statistical methods, including combined principal components analysis (PCA) and linear discriminant analysis (LDA) (10, 42), and soft independent modeling of class analogy (SIMCA) (36), have previously been applied to post-burn ignitable liquid data.

Artificial neural networks (ANN) have been applied to several studies of ignitable liquid classification in the presence of post-burn background contributions (10, 48-51). Classification by ANN has been found to be highly effective for gasoline samples based on near-IR data (49). Harrington and coworkers have applied a fuzzy rule building expert system to differential mobility spectrometry studies of ignitable liquids from fire debris (52, 53). In one example, relevant to the work conducted under this award, LDA was applied to the scores from the first four principal components derived from PCA of the normalized areas of 44 target compounds in 88 samples of two different brands of Canadian gasoline. The premium unleaded and regular unleaded grades were discriminated by LDA at 80-93% depending on the model development and test set composition (10). In another example, LDA was applied to the first three principal component scores resulting from PCA analysis of GC-MS peak area data derived from selected ion monitoring of a set of 35 gasoline samples. The analysis successfully classified 96% of the gasoline samples of the same origin; however, in this study, none of the samples contained substrate pyrolysis products or other sources of interference (42). Results from these studies reflected varying degrees of success when applying the methods to relatively small samples sets.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Classification by LDA and QDA

As part of the research conducted under this award, Bayesian and Fisher linear discriminant analysis (LDA) and Bayesian quadratic discriminant analysis (QDA) were investigated as methods of discriminating between samples that were positive or negative for ignitable liquid residue. The methods were also applied to the task of assigning samples that were positive for ignitable liquid residue into the correct ASTM E1618 classes.

Several books and monograph chapters are available to provide an introduction to the topic of sample classification (54, 55). In a portion of the research conducted under this award, the focus was on two-class classification problems using linear and quadratic discriminant analysis, LDA and QDA, respectively. A brief introduction to LDA and QDA is given here for the benefit of the reader. Bayes LDA and QDA are probability-based hard classification methods that assign class membership based on a classification function, $g(x)$ (55). Hard classification methods require that a sample is assigned to only one class and failure to assign a sample to a class is not an option. The sample is assigned to class ω_i provided that $g_i(x) > g_j(x)$ for classes $i \neq j$, where x is a vector of parameters defining the case in question (54). A minimum error-rate classifier can be achieved by Equation 8, where $p(x|\omega_i)$ is the class-conditional probability density and $P(\omega_i)$ is the prior probability of encountering class ω_i.

$$g_i(x) = \ln[p(x|\omega_i)] + ln[P(\omega_i)] \qquad (8)$$

When the class-conditional probability density is multivariate normal, the discriminant function is given in Equation 9.

$$g_i(x) = -\frac{1}{2}\ln[|\Sigma_i|] - \frac{1}{2}(x-\mu_i)^t {\Sigma_i}^{-1}(x-\mu_i) + \ln[P(\omega_i)] \qquad (9)$$

In Equation 9, μ_i, Σ_i, and $|\Sigma_i|$ are the mean vector, the covariance matrix, and the determinant of the covariance matrix, respectively, for class ω_i. In Equation 9, the term $(x-\mu_i)^t {\Sigma_i}^{-1}(x-\mu_i)$

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

is the square of the Mahalanobis distance from x to the center of class ω_i. Simplifying Equation 9 leads to the following classification function, Equation 10, which is quadratic in x. Equation 10 is general to the multivariate normal case where the covariance matrices are different for each class.

$$g_i(x) = -\frac{1}{2}\ln[|\Sigma_i|] - \frac{1}{2}x^t{\Sigma_i}^{-1}x + x^t{\Sigma_i}^{-1}\mu_i - \frac{1}{2}{\mu_i}^t{\Sigma_i}^{-1}\mu_i + \ln[P(\omega_i)] \tag{10}$$

When the covariance matrices for each class are the same, Equation 10 simplifies to a discriminant function that is linear in x. When the covariance matrices are equivalent across all classes, the $x^t{\Sigma_i}^{-1}x$ term simplifies to $x^t\Sigma^{-1}x$, which is a constant across all classes and can be dropped from Equation 10. The resulting classification function, Equation 11, is now linear in x.

$$g_i(x) = -\frac{1}{2}\ln[|\Sigma_i|] + x^t{\Sigma_i}^{-1}\mu_i - \frac{1}{2}{\mu_i}^t{\Sigma_i}^{-1}\mu_i + \ln[P(\omega_i)] \tag{11}$$

An alternative approach, Fisher discriminant analysis, optimizes the ratio of the between class to within class separation and conveniently produces a set of *n*-1 canonical variates, where *n* is the smaller of the number of classes or number of variables (55). By projecting the multivariate data onto the canonical variate space, separation between the classes may be more readily visualized. Fisher discriminant analysis is utilized in this work to aid in visualizing class discrimination.

Classification by SIMCA

As part of the research conducted under this award, soft independent modeling of class analogy was investigated as method of discriminating between samples that were positive or

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

negative for ignitable liquid residue. The method was also applied to the task of assigning a sample that was positive for ignitable liquid residue into the correct ASTM E1618 class.

SIMCA has been applied to understand complex data in a wide variety of areas including biology (56, 57), biochemistry , fire debris (36), fruit cultivars (58), gasoline (49), meat adulterants (59), oncology (60, 61), pharmaceuticals (62, 63), waste materials (64), and wine (58). One study, relevant to the work undertaken as part of this grant, classified three sample sets of gasoline by "source" and type using the SIMCA method for near infrared (NIR) spectroscopy data. These data sets used by Balabin et al. were based on refinery, process, and type, and correct classification rates of 86%, 70%, and 90%, respectively, were obtained (49). SIMCA has also been used to classify NIR data of tablets used in a clinical study. In this study, De Maesschalck et al. examined the use of the Mahalanobis distance instead of the Euclidean distance for outlier detection. It was found that the modified SIMCA method, utilizing the Mahalanobis distance, yielded better results than the original SIMCA method (62). The Vanden Branden group proposed a robust procedure using both the Mahalanobis and Euclidean distances. For data sets with outliers, this robust approach to SIMCA achieved higher correct classification rates than the original SIMCA method (58).

Soft independent modeling of class analogy is a supervised, soft classification technique which means that the model is built based on a set of training samples with known class assignments and that a sample may be assigned to a single class, multiple classes, or not assigned to any class. The SIMCA method was first introduced in 1975 as a form of pattern recognition that modeled each class by a separate principal components model (65). In this method, a new sample is projected into the principal components space of each class to determine if it belongs to that class (62). The relationship of the projected sample to the class-model PCA space

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

determines if the sample is assigned to that class. One measure of the relationship is referred to as the "orthogonal distance" and represents the Euclidean distance of the sample to the PCA subspace (58). Equation 12 shows the projection of the sample into the PCA model of class *i*, where P_i represents the scores for class *i*. The projected scores, t_i, must then be back-transformed to the original PCA space, Equation 13. Equation 14 shows the calculation of the orthogonal distance for class *i*.

$$t_i = P_i^{\mathrm{T}}(x - \overline{x}_i) \tag{12}$$

$$\hat{x}_i = P_i t_i + \overline{x}_i \tag{13}$$

$$\mathrm{OD}_i = \|x - \hat{x}_i\| \tag{14}$$

A "score distance" was later suggested to improve the original SIMCA method (58). This is a robust version of the Mahalanobis distance and is also measured in the PCA subspace. Equation 15 shows the calculation of the score distance for class *i*, where a_i represents the number of principal components retained and v_{im} represents the eigenvalue for principal component *m* in class *i*.

$$\mathrm{SD}_i = \left[\sum_{m=1}^{a_i} \frac{t_{im}^2}{v_{im}}\right]^{\frac{1}{2}} \tag{15}$$

A linear combination of the scaled orthogonal distance and score distance, referred to as the "score value", is used to classify the new sample, *j*, Equation 16. To scale the orthogonal and score distances, a cutoff value is used for each. The calculations used for the cutoff values for the orthogonal distance, c_{OD_i}, and score distance, c_{SD_i}, are shown in Equations 17 and 18, respectively. In Equation 17, c_{OD_i} is calculated based on the 90% quantile of the standard normal distribution, denoted $z_{0.90}$, and calculates the median and median absolute deviation

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

(MAD) of the orthogonal distance raised to the two-thirds power. In Equation 18, c_{SD_i} is calculated from the 90% quantile of a chi-squared distribution with *a* degrees of freedom (55).

$$d_i^{\mathrm{D}}(j) = \gamma\left(\frac{\mathrm{OD}_i}{c_{\mathrm{OD}_i}}\right) + (1-\gamma)\left(\frac{\mathrm{SD}_i}{c_{\mathrm{SD}_i}}\right) \tag{16}$$

$$c_{\mathrm{OD}_i} = \left[\mathrm{median}\left(\mathrm{OD}_i^{\frac{2}{3}}\right) + \mathrm{MAD}\left(\mathrm{OD}_i^{\frac{2}{3}}\right) z_{0.90}\right]^{\frac{3}{2}} \tag{17}$$

$$c_{\mathrm{SD}_i} = \sqrt{X^2_{a_i,0.90}} \tag{18}$$

Also included in the score value calculation is a tuning parameter, γ, which is a value between zero and one. This can be optimized to give the orthogonal distance and score distance different weights. For example, a γ value of one would base the score value entirely on the orthogonal distance, and a γ value of zero would base it entirely on the score distance. A classification rule can then be made which defines when a sample is assigned to a class based on the score value (55, 58).

In the research conducted under this award, the TIS of fire debris samples were classified using the SIMCA method. Two data sets, one containing only the TIS for samples from the ILRC and Substrate databases (0% substrate contribution) and the other containing the TIS for samples with up to 20% substrate contribution, were used for model development (66). Samples were assigned to the single class that had the minimum score value. A cutoff value was used to determine when a sample would not be assigned to any class. If the minimum score value was above the cutoff value, the sample would not be assigned to any class. The multi-step classification scheme, previously described (66), was also utilized in this research.

Hypothesis or Rationale for the Research

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Current methods of interpreting data from fire debris analysis in the United States relies primarily on ASTM E1618-11, which calls upon the analyst to perform visual pattern recognition aided by comparison with reference materials. This method is currently accepted in court but does not meet the Daubert requirement of a known or potential rate of error, and may be subject to a Daubert challenge in the future. A large number of studies have been conducted that apply various chemometric methods to the analysis of relatively small sets of ignitable liquid and fire debris data. These studies have demonstrated a potential for using chemometric methods for detecting the presence of a specific class of ignitable liquid in laboratory burn samples based on chromatographic data. An alternative approach, using the total ion spectrum (TIS) has been shown to provide specific and differentiable information on complex samples, such as commercial ignitable liquids and pyrolysis products. It is the hypothesis of this research that a combination of chemometric methods and TIS data for a large number of ignitable liquids and pyrolysis products can form the basis for a statistically reliable method of differentiating between fire debris samples that are positive and negative for ignitable liquid residue. The resulting methods will provide the basis for further developments of statistical methods of fire debris analysis that can withstand a future Daubert challenge.

II. Methods

Experimental Methods

Gas chromatography – mass spectrometry data utilized in this study were obtained from three different National Center for Forensic Science sources. Data of reference ignitable liquids were obtained from the Ignitable Liquids Reference Collection database and the data of reference

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

burned substrate materials were obtained from the Substrate database. Data of laboratory generated and large scale fire debris samples were created under a previous NIJ award (2008-DN-BX-K069).

Reference ignitable liquids in the database were typically prepared by diluting 20 µl of the ignitable liquids with 1 mL of carbon disulfide. Samples were analyzed following the protocol previously described (14). Substrate materials were burned by employing a modified destructive distillation method based on a method developed by the State of Florida Bureau of Forensic Fire and Explosives Analysis. Materials were placed upside down in un-lined metal quart paint cans in which the loosely fitted lids had nine 1 mm diameter holes. Heat was applied to the bottom of the cans with a propane torch at a distance of 4 cm. Once smoke appeared, heat continued to be applied for an additional two minutes before the heat was removed. The lids were replaced with intact lids so that vapors would condense while the cans returned to room temperature. The residues from the burned substrate materials were extracted following ASTM E1412-07(68). Prior to sealing the lids, activated carbon strips (10mm x 22mm) were suspended into the headspace of the cans with paperclips and dental floss. The sealed cans were heated for 16 hours at 66°C. Once cooled to room temperature, the activated carbon strips were removed and cut in half. One half was archived and the other was placed into a vial with 500 µL of carbon disulfide for analysis.

Laboratory fire debris samples were created by depositing ignitable liquids, typically 0.5 – 1.0 mL, onto substrate materials prior to heating. Ignitable liquids and substrate materials such as building materials, flooring, and furniture were purchased from local home improvement and furniture stores. Fire debris was created from these materials utilizing the previously described modified destructive distillation method. Residues from the fire debris were extracted following

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ASTM E1412-07 has previously described. Large scale burns were conducted at the Florida Fire College using four 2.4 x 2.4 x 6.1 m^3 Konex shipping containers fabricated inside with sheetrock walls and ceilings to construct a two room structure. Each container had a window and door which allowed limited control of air to support the fire. Flooring, furniture and household items and clothing were placed inside the containers to resemble a bedroom and living room. The containers were re-fitted and re-furnished several times, for a total of twelve fires. Five hundred milliliters of known ignitable liquid was poured throughout the containers by a firefighter who then set a torch to initiate the fire. Duration of the fires was typically 5 – 15 minutes with temperatures reaching as high as 870 °C. Extinguishment of the fires was performed by fire fighters with only water. Upon cool down, 10 – 12 fire debris samples per fire were collected throughout the container and placed into metal paint cans. The residues were extracted from the fire debris following ASTM E1412-07 as previously described.

Statistical Methods

Generation of the Data Sets

Different data sets were generated for each of the statistical methods used in this research. The following sections give descriptions of the data used by each method and how the data sets were generated.

Total Ion Spectra Comparison

The total ion chromatogram (TIC), extracted ion profiles (EIP), and total ion spectra (TIS) compared in this work were produced from Agilent GC-MS data files available in the Ignitable Liquids Reference Collection and Database. The TIC and EIP were generated using the

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ChemStation (version D.02.00.275, Agilent Corp., Santa Clara, CA) software. The TIS were generated from the data matrix exported in the cdf format, by summing the ion intensities over all mass scans at each m/z ratio and normalized so the maximum spectral intensity was equal to one. The data set consisting of 620 IL samples and 88 SUB was obtained from the ILRC and the Substrate Reference Database. The IL samples comprised the following ASTM class (and subclass): 33 GAS, 22 NA, 18 NP, 99 OXY, 154 MISC, 42 ISO, 218 PD (31 LPD, 118 MPD, 69 HPD), and 43 AR. The TIS were generated for the 708 IL and SUB samples in the mass range of *m/z* 30-200, for a total of 171 ions.

In order to compare with the results of Grotch (9), the TIS of the 708 IL and SUB samples are binary encoded to one bit for each mass-to-charge ratio. For each ion in every TIS, if the peak intensity was equal to or greater than a specified cut-off (transition level as a percent of the base peak) it was encoded as a 1; otherwise, the peak was encoded as a 0. Four different transition levels were investigated: 0.0001, 0.001, 0.01, and 0.1.

Cluster Analysis

The data set analyzed for cluster analysis consisted of TIS from 445 IL (436 unweathered IL and 9 weathered gasoline) samples and 88 SUB samples. The TIS were calculated, according to previously published methods with a range of 30-200 *m/z* values (8). The TIS is calculated by summing the intensities of each nominal mass across all retention times (i.e. scan range) with subsequent normalization and removal of any background/baseline contributing ions. This is equivalent to calculating the average mass spectrum across the chromatographic profile. In this

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

work, the major ion found in the baseline, *m*/*z* 32, was eliminated from the data. The TIS were normalized to the base peak which corresponds to the most intense ion.

Self-Organizing Feature Maps

The SOFM training data set consisted of extracted ion spectra (EIS) for liquids in the ILRC database, which were designated as GAS, AR, ISO, NA, NP, LPD, MPD, HPD and OXY ASTM classes. The OXY samples used for training contained predominantly oxygenated compounds and did not have chemical characteristics of other ASTM classes. Classification of the ignitable liquids in the ILRC was performed by a committee of practicing fire debris analysts. The training data set used in this work was comprised of 313 ignitable liquid samples including 289 unweathered ignitable liquid samples as well as 24 weathered GAS samples. The training samples designated as ISO were further sub-classified as light, medium, and heavy (LISO, MISO, and HISO, respectively) based on their carbon range. AR samples were also further sub-classified as light and medium (LAR and MAR) based on their carbon range. The fire debris data set contained 116 EIS of ignitable liquid residues extracted from fire debris samples that were produced in laboratory and large-scale burns. The samples were extracted following the ASTM E1412-07 method (4), and the ignitable liquid residue patterns were observed in the chromatograms. The class designations for the fire debris samples corresponded to the ASTM class of the unweathered ignitable liquid used in the burn. A third data set was compiled, where the samples had chemical characteristics of two ASTM classes. This data set was comprised of 33 MISC and seven OXY samples from the ILRC. Sample preparation and instrumental analysis for all of the data sets are described in previous work (69).

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The EIS for all data sets were comprised of 29 ions chosen as a subset from Table 2 of ASTM E1618-11 (6). These ions represent compound types commonly observed in ignitable liquids within the seven major ASTM classes. Each EIS was normalized to the base peak which corresponds to the most intense ion.

Discriminant Analysis and SIMCA

For the discriminant analysis and SIMCA classification methods, duplicate or triplicate copies of the TIS of the weathered gasoline samples were included in the data set, which resulted in 15 additional IL samples added to this data set. The total of number of IL samples was 460 and there were 88 SUB samples in this data set. Additional samples were created using the total ion spectral data of samples from the ILRC and substrate pyrolysis databases.

An in-house MATLAB (R2011b, Mathworks, Natick, MA) code was used to generate mixed samples containing both IL and SUB contributions. Due to unknown levels of substrate pyrolysis products present in an individual fire debris sample, a data set was created with an upper limit of substrate contribution set at 20%. Testing on data sets containing 50% and 90% substrate contribution yielded similar results by LDA and QDA (66). Each user-generated mixed total ion spectrum contained a percentage of substrate contribution from two randomly selected SUB samples with the remaining percentage of the total ion spectrum composed of contribution from an IL. The total fraction of substrate contribution was randomly selected from a uniform distribution on the interval of [0.01, upper limit]. Using this procedure, a total of 10 IL samples with substrate contribution were generated for each pure IL. Mixed SUB total ion spectra were created using the same method by combining contributions from three substrate pyrolysis total ion spectra. There were approximately five times as many IL samples used from the databases

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

as there were SUB samples. In order to generate approximately the same total number of mixed SUB samples as IL samples, 50 mixed SUB total ion spectra were generated for each individual SUB in the library. The IL samples with substrate contribution were designated as the class of the IL in the sample and the mixed SUB samples were designated as SUB. The total data set contained total ion spectra from 460 pure IL, 4,600 IL samples with substrate contribution, 88 individual SUB, and 4,400 mixed SUB samples.

The data sets with and without substrate contribution included *m/z* 32 and were normalized by dividing the intensity of each *m/z* ratio by the summed intensity value for the sample, which will be referred to as the "summed to one" method. Two additional data sets were created using the samples without substrate contribution. These included *m/z* 32 and were normalized by the "base peak" and "unit vector" methods. The "base peak" method is calculated by dividing the intensity of each *m/z* ratio by the intensity for the base peak of that sample. With this method, the intensity for each *m/z* ratio is represented as a percentage of the base peak intensity. The "unit vector" method involves dividing the intensity of each m/z ratio by the square root of the sum of squares for the sample.

Fire Debris Samples

Laboratory fire debris samples were generated following the procedure previously described (66). In this method, a specific volume of IL was added to a substrate or multiple substrates placed in an unlined paint can, covered with a vented lid, and placed above the flame of a propane torch. Once smoke was observed to exit the can through the vent holes, the can was kept over the flame for an additional two minutes. After removal from the flame, the vented lid was replaced with a solid lid, and the can was cooled. The ignitable liquid residue was extracted

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

following the ASTM E1412-12 standard for adsorption onto activated charcoal strips (3). The GC–MS method utilized for analysis of the samples has been previously described (8). Substrate control samples were created following this same procedure, but without the presence of an IL.

Large-scale burns were conducted in freight containers staged as a furnished two-room apartment (70, 71). An IL of known ASTM E1618-10 class was used in each container and the location of the IL pour was recorded. Fire debris samples were collected from various areas of the container both on and off the IL pour. Two approaches were taken to assigning the "ground-truth" classes to the fire debris samples. In one approach, laboratory-burn samples were designated as the class of IL used in the burn, and the substrate control samples were designated as SUB. The results using this method will be referred to as "fire debris – pour." Using this approach for the large-scale burns, samples were designated as the class of IL used in the burn if they were collected from the pour, and designated as SUB if they were collected off the pour. The laboratory-burn data set consisted of 69 samples designated as IL and five multiple-substrate control samples designated as SUB. The large-scale burn data set contained 89 samples designated as IL and 40 designated as SUB. Sample designation using this approach considers the pre-burn sample composition and may not account for the loss of IL during the fire. The second approach to assigning the ground-truth classes to fire debris samples, takes into account these potential losses. In the second approach, the data of both the laboratory and large-scale fire debris samples were examined by an analyst[1] with prior knowledge of the ignitable liquid used in the burn. It should also be noted that the analyst had access to the TIC for the unevaporated liquid. Following ASTM E1618-10 protocols (72), the analyst determined whether or not an IL residue was present in each sample. If evidence of IL residue was observed, the sample was

[1] Co-author, M. Williams,holds an MS in Forensic Science, has developed a large online database of ignitable liquid forensic data over the past 12 years, and works closely with the SWGFEX Ignitable Liquids Reference Collection and Database committee to analyze and classify database entries according to ASTM E1618.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

designated as the class of IL used in that burn. If no evidence was observed, the sample was designated as SUB. It should be noted that validation of the statistical models must be made in comparison to the "ground truth", which requires a definitive knowledge of whether samples contain ignitable liquid residue. Validation is not achieved by comparison to the performance of a trained fire debris analyst. The results using this method will be referred to as "fire debris – analyst." Using this approach for the laboratory burns, 69 samples were designated at IL and five samples were designated as SUB. It should be noted all laboratory burn samples that had an IL added, also had evidence of IL residue. For the large-scale burns, this approach resulted in 54 samples designated as IL and 75 designated as SUB. Prior knowledge of the IL chromatographic patterns and target compounds makes data analysis less subjective and more accurate.

The designations of the laboratory burn samples did not change when considering the analyst's examination of the data instead of the pre-burn conditions. Since the designations of the large-scale burn samples were affected by these approaches, the fire debris data will be described as having designations based on the proximity to the pour or based on the analyst's examination of the data. The total fire debris data set consisted of the TIS for 158 samples designated as IL and 45 samples designated as SUB, for a total of 203 fire debris samples, when using designations based on the proximity to the pour. The analyst's examination of the data resulted in 123 samples designated as IL and 80 samples designated as SUB. Correct classification rates using both methods of assigning the ground truth will be discussed.

The TIS of the fire debris samples were normalized using the same methods that were employed for the model data sets.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Computational Details

Total Ion Spectra Comparisons

Total ion spectra comparison calculations were performed in RStudio version 0.97.312 (RStudio Inc., Boston, MA) using code written in-house. Comparisons of binary encoded spectra were made on a pairwise basis for all unique TIS pairs. The number of disagreeing ions was recorded for each comparison and the average number of disagreements and the distribution of disagreements were calculated.

As the ability to discriminate between ion intensities increases, the number of disagreeing ions in a spectral comparison will increase and the number of failures to discriminate in a series of comparisons will decrease. In order to investigate the magnitude of this trend, discriminations and the number of disagreeing ions were determined for comparisons of TIS that had not been binary encoded. A 0.01 transition level for peak detection was implemented and two ions were considered to exhibit a disagreement if the relative standard deviation (RSD) for a comparison was greater than or equal to a specified limit. For each pairwise comparison of ion intensities, the relative standard deviation was calculated as the ratio of the population standard deviation to the average of the two intensities of the ion being compared. The population standard deviation is used because in each comparison (i.e., comparison of m/z 91 in TIS-1with m/z 91 in TIS-2) there are only two ions to compare. Comparison of an intensity of 0 and an intensity of 1, as for binary encoded spectra, corresponds to a RSD of 1 (i.e., a %RSD of 100). The number of disagreeing ions in a spectral comparison and the number of discriminations between spectra were calculated at a series of relative standard deviations.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Cluster Analysis

Hierarchical cluster analysis was performed using a variety of distance and linkage combinations where the cophenetic correlation coefficient was calculated for each combination. The cophenetic correlation coefficient corresponds to how well the resulting clustered data matrix reflects the original data matrix. The clustering solution most accurately reflects the original data when the cophenetic correlation coefficient is equal to one (20). The distance-linkage combination with the largest cophenetic correlation coefficient and chemically meaningful groups was chosen for further analysis. Optimal leaf ordering was subsequently performed which "maximizes the sum of the similarities of adjacent leaves" in the dendrogram, to allow samples that are highly similar to be arranged in the center of a cluster (73). The leaves on the dendrogram correspond to individual samples. Calculations were performed in R (74) with the following packages available on the Comprehensive R Archive Network (CRAN): the amap package (75) for the distance calculations and the cba package for optimally ordering the dendrogram leaves. The stats package (74), used for hierarchical clustering and cophenetic correlation coefficient calculations, is normally loaded as part of the basic R software.

Similarities were calculated from the distance metric selected for clustering and a similarity matrix, consisting of the similarity values for each pairwise comparison of the samples, was generated. The similarity matrix was exported in a comma-separated format and read into Excel (Microsoft Corporation, Redmond, WA) for further analysis. A heat map can be created by assigning a color gradient to the range of similarity values within the similarity matrix. The color gradient was assigned using the conditional formatting option in Excel. The selected condition applied white to the cells with the lowest similarity values, gray to cells at the 75th percentile, and black to the cells with the highest similarity values.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Figure 1 shows an example heat map created from the similarity matrix. In Figure 1, pairwise comparisons were made between samples to create the similarity matrix. The lowest similarity value was 0.454 and shaded white, while the values with the highest similarity (1.000) are along the diagonal and shaded black. Cells with values between 0.454 and 1.000 are various shades of gray. From the example in Figure 1, samples labeled SRN 320 and 298 are easily observed to be highly similar to one another, as well as SRN 46 to 107.

SRN	320	298	400	27	46	107	385	75
320		0.977	0.873	0.850	0.894	0.925	0.702	0.474
298	0.977		0.869	0.850	0.590	0.925	0.692	0.473
400	0.873	0.869		0.925	0.955	0.940	0.671	0.454
27	0.850	0.850	0.925		0.958	0.917	0.703	0.468
46	0.894	0.590	0.955	0.958		0.977	0.795	0.548
107	0.925	0.925	0.940	0.917	0.977		0.810	0.596
385	0.702	0.692	0.671	0.703	0.795	0.810		0.738
75	0.474	0.473	0.454	0.468	0.548	0.596	0.738	

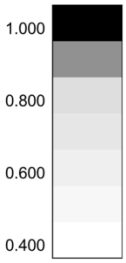


Figure 1: Generation of a heat map by applying conditional formatting to a similarity matrix. SRN represents the sample reference number used to identify the ignitable liquid in the ILRC database. The lowest similarity values (0.4) are shaded white, the highest similarity values (1.0) are shaded black, and gray is used for the midpoint values. A color gradient was applied to the cells with similarity values between 0.4 and 1.0.

Self-Organizing Feature Maps (SOFM)

SOFM calculations were performed using the Neural Network Toolbox™ 7 with MATLAB R2011b (MathWorks, Natick, MA, USA). Several hexagonal grid sizes were investigated. The 15x15 unit hexagonal grid resulted in the minimal number of neurons being

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

associated with more than one ASTM class. A total of 112,500 epochs, 500 times the number of neurons were calculated, where the ordering phase of learning consisted of 1,000 epochs and an initial neighborhood size of 13 (26). Therefore, in the first epoch of the ordering phase, the winning neuron and the neurons within a 13 neuron radius from the winning neuron had their weight vectors adjusted. The neighborhood size decreased from 13 to one as the number of epochs in the ordering approached 1,000. The subsequent tuning phase had a neighborhood size of one, so only the weight vector of the winning neuron was adjusted. The default batch training method in the Matlab Neural Network Toolbox™ 7 was used, where for each epoch, the whole data set was presented to the network. The winning neuron for each input vector was determined, and the weight vectors were moved to the average position of all input vectors for which the weight vector was a winner or in the neighborhood of a winner (30). In the batch training mode, a learning-rate parameter was not used (25).

The EIS for the fire debris data set were projected onto the trained SOFM by calculating the distances between the final weight vector for each neuron and the variable vector of each sample. The sample was assigned to the neuron for which its weight vector had the smallest distance. The assigned neuron may have been associated with a training sample or may be an empty neuron (55). Select ignitable liquid samples from the ILRC database designated to the OXY class or MISC category were also projected onto the SOFM.

Discriminant Analysis Methods

MATLAB was used to perform PCA, LDA, QDA, and SIMCA. To reduce the dimensionality of the data, PCA, a multivariate analysis method, was performed on each data set comprised of the total ion spectrum from each sample. Factorization of the data matrix was accomplished by singular value decomposition. The rank of the data matrix was reduced based

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

on the number of principal components retained to reproduce 50%, 70%, 90%, or 95% of variance in the data. The number of principal components retained was varied in order to optimize the percentage of samples classified correctly. The scores for the retained principal components were used for the development of linear and quadratic discriminant functions.

In the application of LDA and QDA to the data, the assumption is made that the prior probabilities of encountering each class are equal. The problem is reduced to a set of two-class comparisons and the prior probability of each class is set to 0.5. Scores, derived from the PCA of the set of total ion spectra for the two classes being modeled, were used as variables for the development of the LDA and QDA models. The fraction of variance retained in the final LDA or QDA model was based on the optimized correct classification rate. Discriminant functions, given by either the quadratic or linear combination of the PCA scores for each class, were based on Equations 10 and 11, respectively. The sample's classification is predicted by the class which has the largest quadratic or linear discriminant score (55).

Multi-Step Classification Procedure

Classification by LDA or QDA could theoretically be implemented wherein a sample is assigned to one of the ASTM classes or the SUB class in a single step; however, in practice, a problem arises with implementing the QDA method. The QDA model requires a larger number of adjustable parameters than LDA, due to the larger number of covariance matrices (i.e., the covariance matrices for all classes are not assumed equal), which can lead to a small sample size problem (76). The small sample size problem arises when the number of training samples is less than or comparable to the sample space dimensionality. This problem was encountered for some of the ASTM classes with smaller representation in the training data set (i.e., NA and NP). The

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

small sample size problem was obviated by a multi-step classification approach where two classes were considered in each step.

In the multi-step classification procedure, the first step discriminates between samples that contain an ignitable liquid residue and those that only contain substrate pyrolysis products, as depicted in Figure 2. In the second step, samples designated as containing ignitable liquid residue are further classified as containing residues from the combined AR and GAS class (denoted AR/GAS) or as one of the remaining ASTM classes that contain predominantly aliphatic (denoted ALI) compounds. In the next level of discrimination, as reflected in Figure 2, the samples designated as AR/GAS are further classified as AR or GAS, and ALI samples are further classified as containing ignitable liquid residue from the ASTM classes in the ISO/NA set or the ASTM classes in the PD/NP set. Samples designated as either the PD/NP set or ISO/NA set are further classified as one of the constituent ASTM classes. Finally, the data for fire debris samples classified into the AR, GAS, PD, NP, ISO, or NA classes should be examined to determine if they should be reclassified into the OXY class or MISC category. Under ASTM E1618-10, an ignitable liquid from any of the other classes can be assigned to the OXY class if it contains a significant oxygenated component as indicated by the major ions, 31 and 45 m/z (i.e., a PD containing methyl ethyl ketone would be classified as an OXY). Notably, the oxygenated components are frequently low boiling and easily lost due to weathering. The miscellaneous category does not possess class characteristics and, therefore, is a “catch-all” designation of last resort for ignitable liquids that do not belong to the other ASTM E1618-defined classes. Because LDA and QDA are hard classification techniques, samples that contain mixed ignitable liquids will be classified as the ASTM class the spectral data most resembles; however, the

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

correct assignment for this type of sample is actually to the miscellaneous category since it is composed of a mixture of ignitable liquids.

In the model development for the multi-step classification procedure, samples used in the first step (IL/SUB discrimination) are from all seven IL classes and the MISC category. In all subsequent steps, the samples designated as OXY or MISC samples are removed from the model data sets.

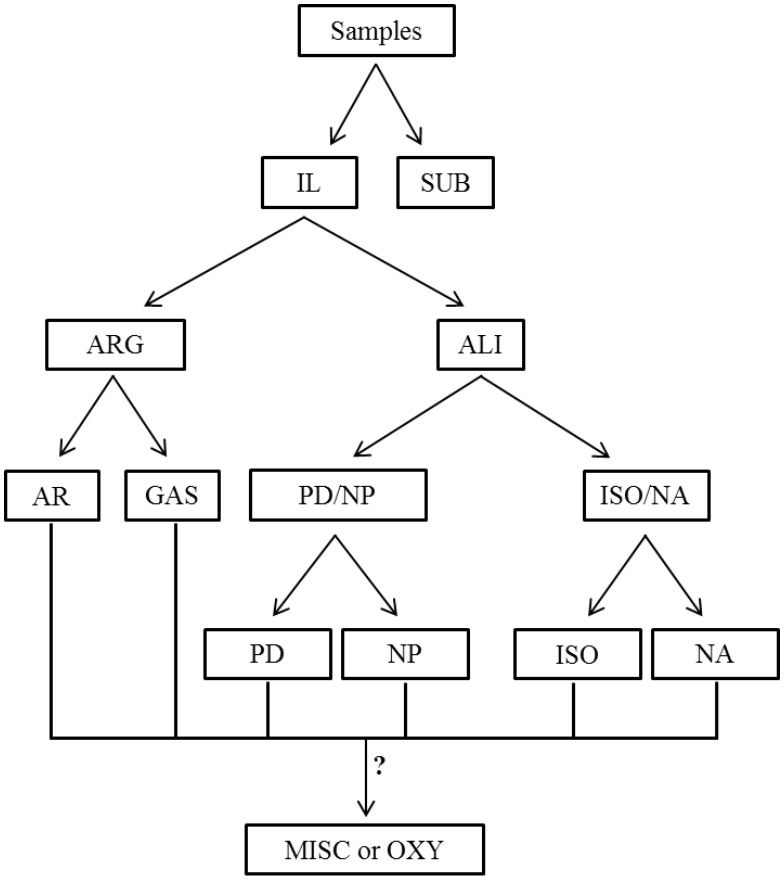


Figure 2: Multi-step classification scheme used for model development. Correct classification rates for each step of this classification scheme are independent of the previous step.

To determine the error rate for each step of the multi-step classification procedure, a MATLAB code that applied cross-validation and discriminant analysis to the data were written in-house. On each of the 100 iterations conducted, 20% of model data were randomly selected

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

and held out as a test set and the remaining 80% was used for LDA or QDA model development. The correct classification rates were based on the classifications of the test set.

A fire debris data set, described above, was used to test the LDA and QDA models. The fire debris data set was projected into the model PCA space and the projected scores were used to classify the fire debris samples by the LDA and QDA models.

Likelihood Ratio Test

As previously discussed, LDA is a simplified case of QDA in which the covariance matrices are assumed to be equivalent. If the covariance matrices are determined to be statistically different, QDA is the more appropriate method of analysis (77). The likelihood ratio was used to compare the covariance matrices for the two classes used in each step of the classification scheme (78).

Evaluation of Classifier Performance

For the two-class comparison of ignitable liquids or ignitable liquid residue-containing samples (collectively IL) and substrate-only samples (SUB), the false positive, true positive, false negative, and true negative rates were calculated for each of the data sets. The false positive rate is defined as SUB samples classified as IL. The true positive rate is defined as IL samples classified as IL (79). In the forensic application, a low false positive rate is important. The false negative rate represents IL samples that were incorrectly classified as SUB, while the true negative rate represents SUB samples that were correctly classified as SUB. The false negative rate can be calculated as one minus the true positive rate. Similarly, the true negative rate can be calculated as one minus the false positive rate.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

SIMCA

Prior to model development, the PCA scores were used to identify outliers and leverage points within the data. PCA was performed on the data set for each individual class in the given step of the classification scheme (i.e. samples designated as IL were separated from those designated as SUB prior to performing PCA). The orthogonal distance, Equation 14, and score distance, Equation 15, were calculated for each sample in each class. These values were then compared to the cutoff values for the orthogonal distance and score distance calculated for each class, Equations 17 and 18, respectively. Orthogonal outliers, defined as having an orthogonal distance greater than c_{OD_i}, but a score distance less than c_{SD_i}, were excluded from the data sets. Bad leverage points, defined as having an orthogonal distance greater than c_{OD_i} and a score distance greater than c_{SD_i}, were also removed from the data sets. These points are far from the class mean when projected into the PCA space. Good leverage points were retained for model development and stabilize the estimation of the PCA space. Good leverage points have a score distance greater than c_{SD_i}, but an orthogonal distance less than c_{OD_i}. Samples with an orthogonal distance and score distance below the respective cutoff values were also used for model development (55, 58).

MATLAB code was written in-house to perform outlier removal, PCA, and the SIMCA method with cross-validation. For cross-validation, a model was developed for each class in the given step of the classification scheme. Orthogonal outliers and bad leverage points were removed from the data set prior to model development. On each of the 100 cross-validation iterations, 20% of model data were randomly selected and held out as test samples, while the remaining 80% was used for model development. The data used for model development was

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

analyzed by PCA, and the scores and eigenvalues for the retained principal components were used in the calculation of "score values" ($d_i^{\mathrm{D}}(j)$ from Equation 16) for sample classification. The TIS for test samples were assigned to the class with the minimum score value; however, if the minimum was greater than a value of one, they were not assigned to a class. Test samples were not assigned to multiple classes. Correct classification rates for cross-validation were based on the classifications of the test set.

The fire debris data set, described previously, was used to test the models. The TIS of fire debris samples were projected into the model PCA space and the projected scores were used for classification. The TIS of the sample was assigned to the class with the minimum score value; however, if the minimum was greater than a value of one, it was not assigned to a class. Samples were not assigned to multiple classes. The value of gamma was optimized for each step of the classification scheme and the optimal value was selected based on the highest correct classification rates for the cross-validation test samples.

Classifier performance metrics were also calculated for the results obtained using SIMCA.

Training

The QDA method, described above, was trained to two groups of fire debris analysts by both online and face-to-face training. The online and face-to-face training required a maximum of eight hours to complete. The online material was delivered using a popular learning management system and students had one week to complete the training. The face-to-face material was delivered over an eight hour period in a single day. A summative assessment was performed to evaluate student comprehension. The summative assessment consisted of a multiple choice exam composed of questions that examined the student's understanding of the theoretical basis for the statistical methods and their ability to complete a practical examination of data from real fire

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

debris using the software provided during the training. An outline of the training content is given in Appendix I at the end of this report.

III. Results

Comparison of Total Ion Spectra

Figure 3 shows the comparisons made between the TICs and TIS for the GAS, NA, PD and MISC IL samples analyzed on the same in-house instrument. Consider the TICs and TIS corresponding to the GAS and NA classes; while the TICs are visually quite different within a class, the TIS are very similar. However, more similar patterns of ions are observed when comparing the TIS for samples in the NA and ISO class (not shown). Large variations in both the TICs and TIS are observed for samples within the MISC category, which is not unexpected given that these samples may have characteristics of more than one ASTM class.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

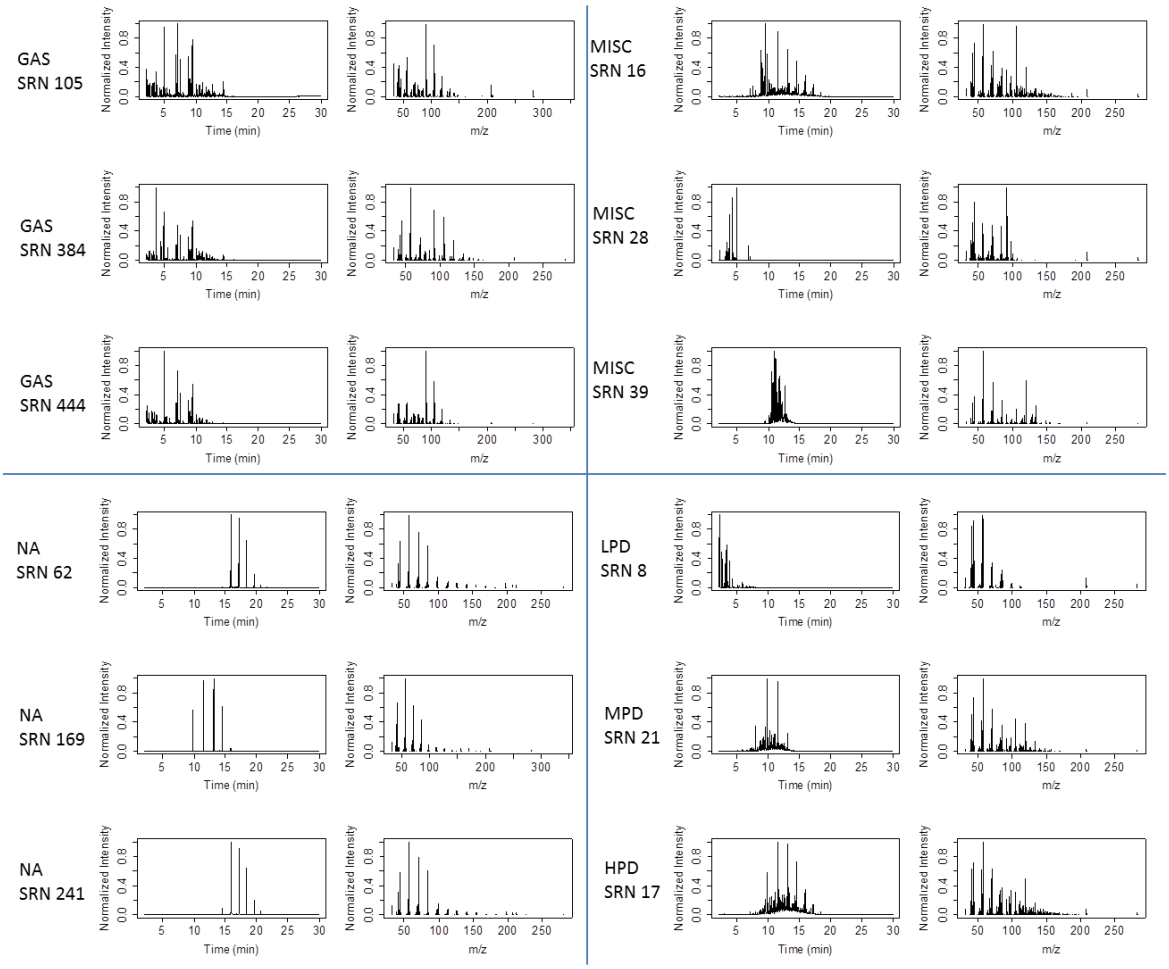


Figure 3: IL TIC (left) and TIS (right) Grouped by ASTM Class

The TIS for the 708 IL and SUB samples were binary encoded based on the four different transition levels. The information content for transition levels 0.0001, 0.001, 0.01, and 0.1 were calculated, using Equation 7, to be: 114.2, 125.1, 121.9, and 56.53 bits, respectively. These results are consistent with those found by Grotch for approximately 3,000 organic compounds (9). The number of disagreeing ions between each of the 250,278 unique pairwise comparisons of the 708 IL and SUB samples was determined, and the frequencies are shown for each of the transition levels in Figure 4. The average observed number of disagreeing ions for transition

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

levels 0.0001, 0.001, 0.01, and 0.1 are: 52.61, 58.16, 56.08, and 23.69 ions, respectively. The results at the one percent (0.01) transition level are consistent with previous analyses of 440 IL samples binary encoded at the same level for TIS with *m/z* 30-350, which were found to have an average of 50 differing ions per comparisons for 96,580 unique pairwise comparisons (8). The notable drop in the number of disagreeing ions and information content as the transition level is increased from 0.01 to 0.1 indicates a significant portion of the discriminating ions are within this relative intensity range. Out of the 250,278 unique pairwise comparisons of the 708 IL and SUB samples at the 0.01 transition level, only 20 comparisons resulted in zero ion disagreements (approximately 8×10^{-3} %, which extrapolates to approximately 80 indistinguishable spectra in 10^6 comparisons). Grotch found six indistinguishable spectra in 10^6 comparisons for pure organic compounds (9). Each of the 20 indistinguishable comparisons in this work occurred between IL samples with the same ASTM class (or subclass). One of these pairwise comparisons occurred between two NP samples and another between two OXY samples. Three indistinguishable pairwise comparisons occurred between MISC samples, one comparison for LPD samples, 10 comparisons for samples designated as MPD, three comparisons for HAR samples, and one comparison between two MAR samples. Grotch found that approximately three perfectly matched spectra occurred per 10^6 comparisons of low resolution mass spectra of organic compounds, where the matched spectra corresponded to isomers. These results demonstrate that even when binary encoded at the transition levels studied in this work, the "mass spectrum is a very specific chemical signature" (9), and while 80 indistinguishable pairs of spectra for each 10^6 comparisons is significantly larger than three indistinguishable pairs, the binary encoded TIS comprise a significantly unique signature for complex mixtures, such as ignitable liquids.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

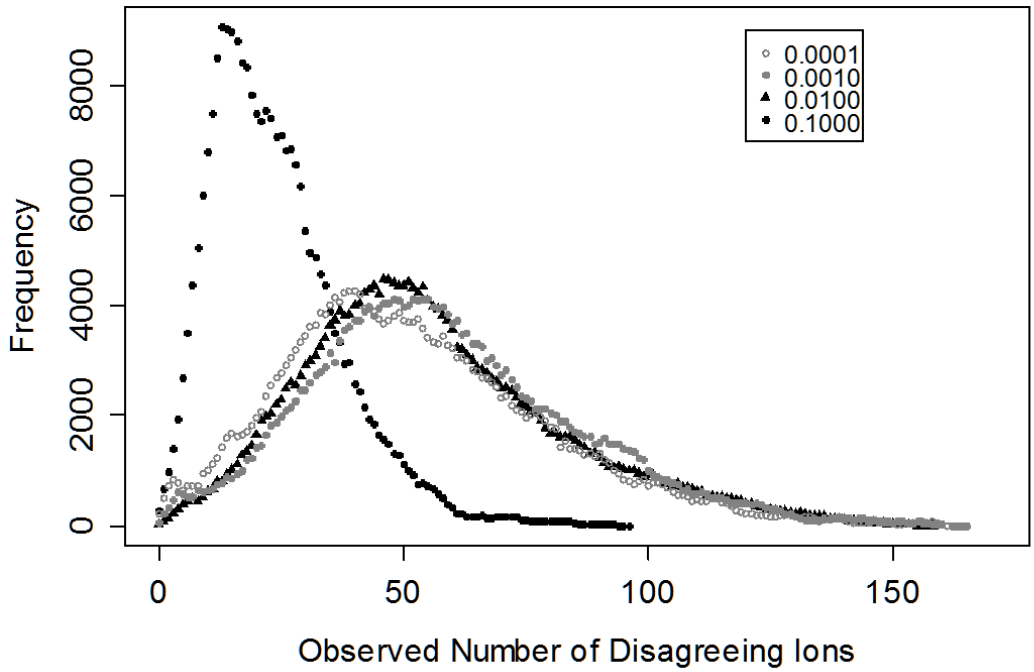


Figure 4: Frequency of Number of Disagreeing Ions for 708 IL and SUB Samples for Varying Transition Levels.

The average observed numbers of disagreeing ions in pairwise comparisons of binary encoded TIS were then calculated, at the one percent transition level, for pairwise comparisons of samples with the same and differing class designations. The resulting averages are shown as a matrix in Table 1, where smaller averages are shaded lighter and larger averages have darker shading. Pairwise comparisons of samples with the same designation are given along the diagonal, while pairwise comparisons of samples with different designations are given off the diagonal. For the 620 IL-designated samples, an average of 54.6 disagreeing ions was found for 191,890 pairwise comparisons. Similarly, the 88 SUB samples with 3,828 unique pairwise comparisons had an average of 56.0 disagreeing ions. When considering the 54,560 unique pairwise comparisons that occur between samples designated as IL and SUB, an average of 61.3

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

disagreeing ions was observed. The largest average number of disagreeing ions in pairwise comparisons was between the HPD and OXY classes (98.6), which reflect the significant difference in chemical composition and molecular weight range between these two classes. The smallest average number of disagreeing ions was found between ISO and NA classes (16.8), which reflect the similar MS fragmentation behavior of the chemical components of these two classes.

Table 1: Average number of differing ions for comparisons of binary encoded TIS between differing ASTM ignitable liquid classes.

	IL	GAS	NA	NP	OXY	MISC	ISO	LPD	MPD	HPD	AR	SUB
IL	54.6											
GAS		25.3										
NA		54.3	8.8									
NP		70.3	55.0	23.4								
OXY		53.5	46.7	75.2	47.3							
MISC		49.3	51.0	54.6	59.3	52.0						
ISO		51.6	16.8	57.3	43.3	50.7	14.2					
LPD		46.7	27.9	55.5	47.1	47.7	28.9	25.7				
MPD		48.5	38.7	38.6	59.1	44.6	40.8	38.8	22.5			
HPD		78.6	85.4	44.3	98.6	70.4	89.8	82.7	59.5	30.3		
AR		44.5	63.7	84.5	52.9	61.3	57.6	58.5	65.6	97.7	34.8	
SUB	61.3	52.6	66.3	60.5	67.7	58.7	66.1	60.4	53.7	68.4	66.7	56.0

The results given here compare favorably with those presented by Grotch (9), but show that the binary encoded TIS to lead to an extrapolated larger number of failures to discriminate in 10^6 comparisons of complex mixtures, such as ignitable liquids. When the TIS are encoded to a higher degree of accuracy, the average number of disagreeing ions per pairwise comparison is expected to increase and the numbers of failures to discriminate are expected to decline. To test this assertion and to gain a better estimate of the performance of the TIS, a set of pairwise comparisons was made of spectra that had not been binary encoded. Disagreeing ions were determined at RSD of 0.05, 0.1, 0.2 and 0.4, as described in the Materials and Methods section. The average numbers of disagreeing ions per comparison were found to be 92, 89, 83 and 74

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

respectively. Recall that 56 disagreeing ions were found for binary encoded TIS with a 0.1 transition level, which corresponds to an RSD of 1. These results are shown graphically in Figure 5. A set of five repeat injections of a standard mix of aliphatic and aromatic hydrocarbons was determined to have an RSD of 0.1 for the TIS generated with a 0.01 peak detection transition level. The TIS comparison at an RSD of 0.1 resulted in three failures to discriminate in 250,278 comparisons. Two of the failures to discriminate were between ignitable liquids of the MPD class and the remaining failure to discriminate was between ignitable liquids of the HPD class.

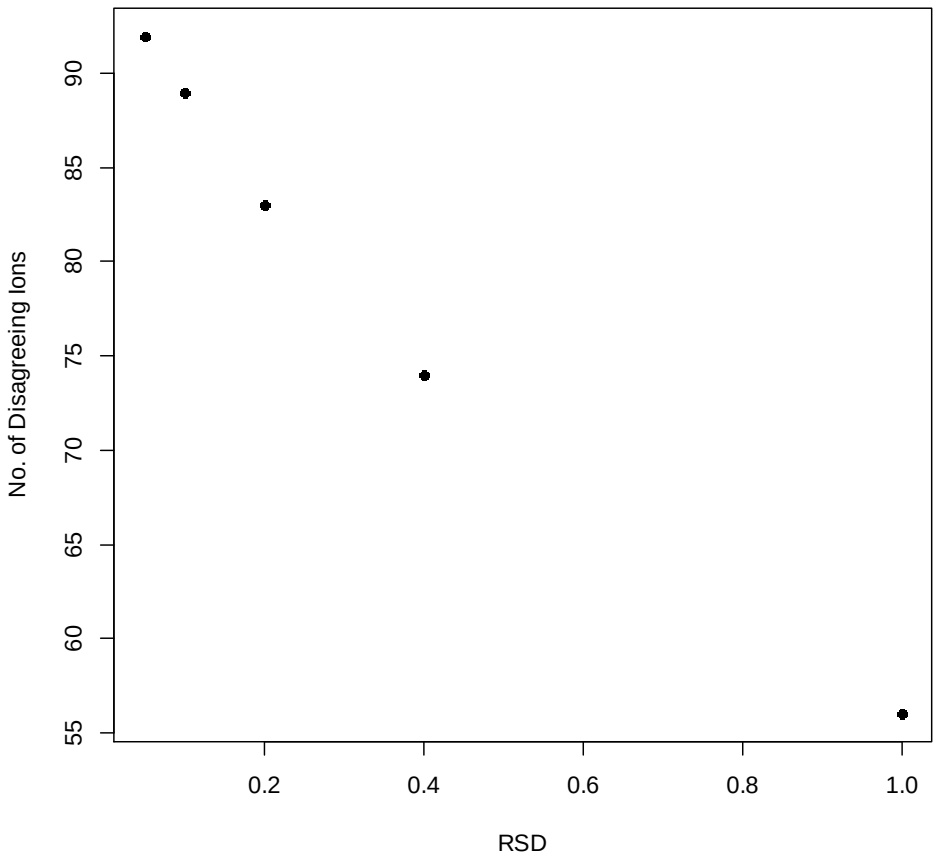


Figure 5: The average number of disagreeing ions per TIS comparison is plotted versus the relative standard deviation cutoff for TIS having a 0.01 peak transition level. The RSD=1 point corresponds to the comparison of binary encoded spectra at a 0.01 peak transition level.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Grouping of Ignitable Liquids

Cluster Analysis

The correlation distance, *d*, was selected for this research and is equivalent to 1-*r*, where *r* is the Pearson correlation coefficient. The Pearson correlation coefficient ranges from -1 to 1; therefore, the distance values range from 0 to 2. Similarity values were calculated by 1-(*d*/2) and can range from 0 to 1. Samples with the smallest distance will, therefore, have the largest similarity. The similarity matrix for 445 IL samples and 88 SUB samples was converted to a heat map, as previously described, and is shown in Figure 6. The samples were arranged based on ASTM class designations found in the ILRC database . Within each class, the samples were arbitrarily ordered; therefore, the most similar samples may not be adjacent to one another.

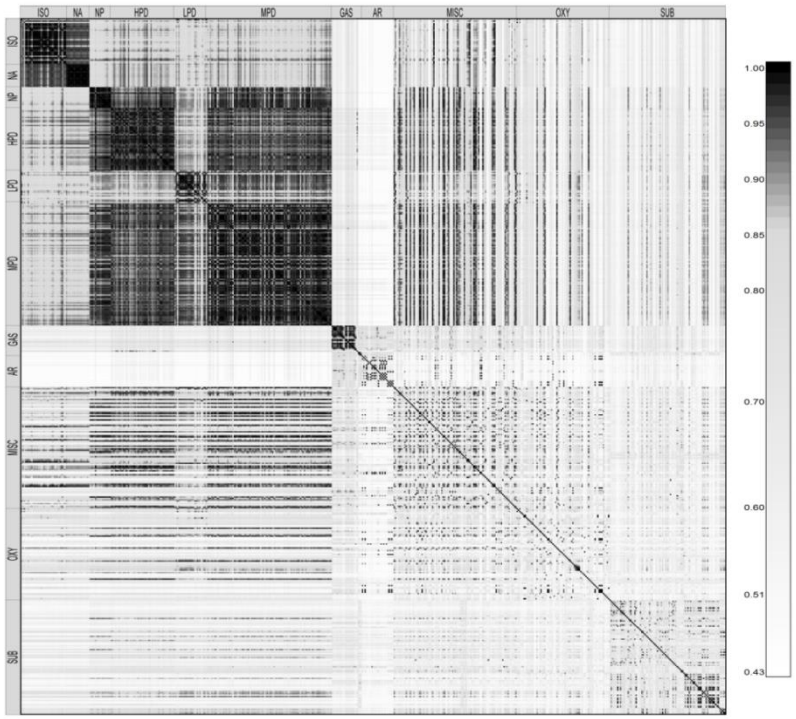


Figure 6: A heat map for all 445 IL and 88 SUB samples. The samples are sorted by ASTM classes but are arbitrarily ordered within the classes.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

A high degree of similarity is observed between IL within some ASTM classes, but not within others. Variation in the similarity values within the SUB, MISC, and OXY groups is reflected in the shading of the lower right corner of Figure 6. Samples in the MISC category and OXY class are expected to have significant variation based on the criteria for assignment to these classes, outlined in ASTM E1618-11. When ILs do not meet the criteria of an ASTM class or when they possess characteristics of multiple ASTM classes, they are assigned to the MISC category. Assignment to the OXY class occurs when the IL contains a significant oxygenated component with an abundance at least one magnitude greater than the matrix peaks of the chromatogram and mixtures of oxygenated compounds and other compounds may be present (6). The remaining ASTM classes (ISO, NA, NP, PD, GAS, and AR) possess characteristics which allow for better defined classification criteria.

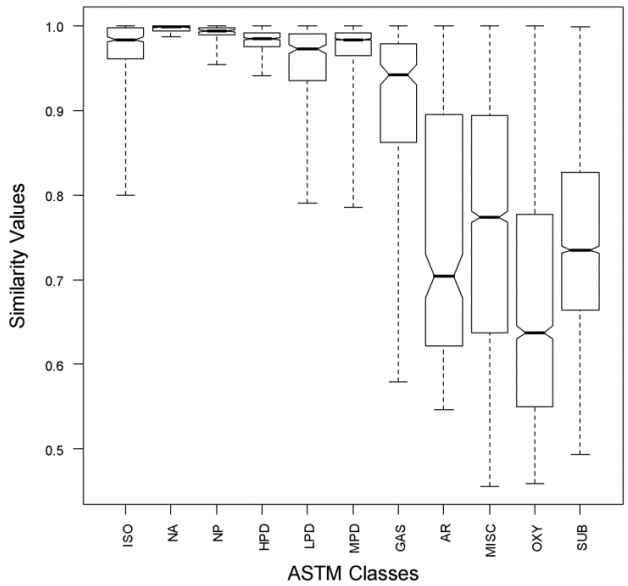


Figure 7: Box plots for the similarities within each ASTM class. The horizontal lines outside of the box represent the minimum and maximum similarity values. The median similarity value is indicated by the thick, dark horizontal line and the box is constructed using the 25th and 75th quartiles. The notch represents the confidence interval around the median (± (1.57 × "IQR")/√n, where IQR is the interquartile range and n is the number of samples).

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The variation of similarity values within the ASTM classes is demonstrated by the box plots in Figure 7. The similarity between TIS within an ASTM class represents consistency in chemical composition and MS fragmentation patterns. For example, the box plot for NA samples in Figure 7 shows a small range and high median value of similarities, which reflect a high consistency in chemical composition and MS fragmentation patterns. The NA class contains liquids with normal alkanes in a high carbon range. The AR class exhibits a broader variance and lower median similarity value, which reflect the variation in MS fragmentation between alkylbenzenes and alkylnaphthalenes, and the variation in chemical composition within the class.

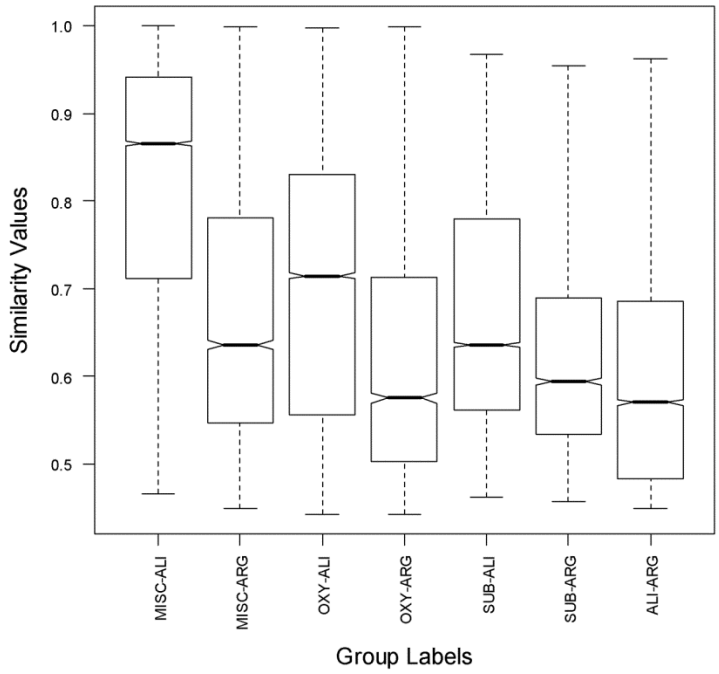


Figure 8: Box plots for between-class similarity comparisons. The horizontal lines outside of the box represent the minimum and maximum similarity values. The median similarity value is indicated by the thick, dark horizontal line and the box is constructed using the 25th and 75th quartiles. The notch represents the confidence interval around the median (± (1.57 × "IQR")/√n, where IQR is the interquartile range and n is the number of samples).

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Similarity values between groups of classes are demonstrated by the box plots in Figure 8. Classes in group ALI are ISO, NA, NP, HPD, LPD, and MPD because these classes predominantly contain aliphatic compounds. Classes in group ARG are GAS and AR, because these classes predominantly contain aromatic compounds. In Figure 8, MISC, OXY, and SUB samples are observed to be more similar to the aliphatic group (ISO, NA, NP, and PD classes) than to the aromatic group (GAS and AR classes). Comparison of the aliphatic group to the MISC, OXY, and SUB classes resulted in decreasing median similarity values, respectively. The low median similarity value between the aliphatic and aromatic groups indicates the difference in chemical composition of samples within these groups.

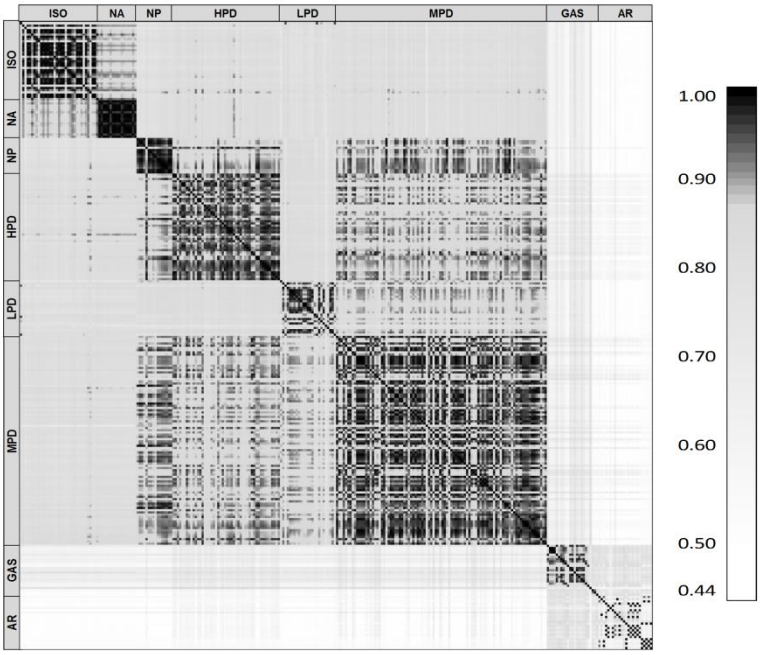


Figure 9: A heat map using only those 282 IL samples classified as ISO, NA, NP, PD, GAS, or AR. The samples are sorted by ASTM class, but are arbitrarily ordered within the classes.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Classes that have the least within-class variation (ISO, NA, NP, PD, GAS, and AR) in Figure 7 are well-defined by ASTM and were considered for further analysis (note that subclasses of AR have more characteristics TIS than other classes, see discussion on pg 55). The 282 samples that comprise these classes were re-plotted in Figure 9 and remain in the same order as Figure 6. As can be seen in Figure 9, there are two distinct groups: aliphatic and aromatic. The darkest shading reflects same sample comparisons which can be seen along the diagonal of the heat map. Shading around the diagonal indicates higher similarity values for samples within the same class compared to lighter shading of the between-class samples, shown on the off-diagonal.

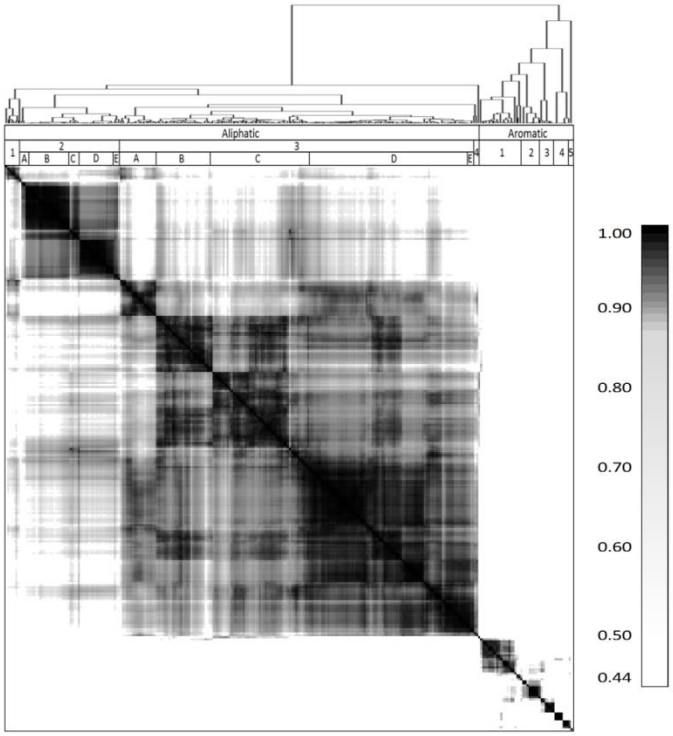


Figure 10: A dendrogram and associated heat map for the 282 IL samples classified as ISO, NA, NP, PD, GAS, or AR. The samples are sorted based on the order of the dendrogram obtained from cluster analysis.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Clustering was performed on the same data set used in Figure 9. The heat map and resulting dendrogram, obtained using the correlation distance and average linkage for cluster analysis, is shown in Figure 10. In this heat map, the samples were reordered based on the clustering results. The sample order is the same for the heat map and the dendrogram in Figure 10 and the location of the samples in the heat map is directly below their location in the dendrogram. The color gradient, which illustrates the degree of similarity, can be observed for the clustered samples. The reordered heat map and the dendrogram indicate two distinct clusters which correspond to the aliphatic and aromatic groups, as previously described. These clusters will be referred to as the Aliphatic and Aromatic Clusters. Labels were assigned to the dendrogram to assist in discussing the clusters.

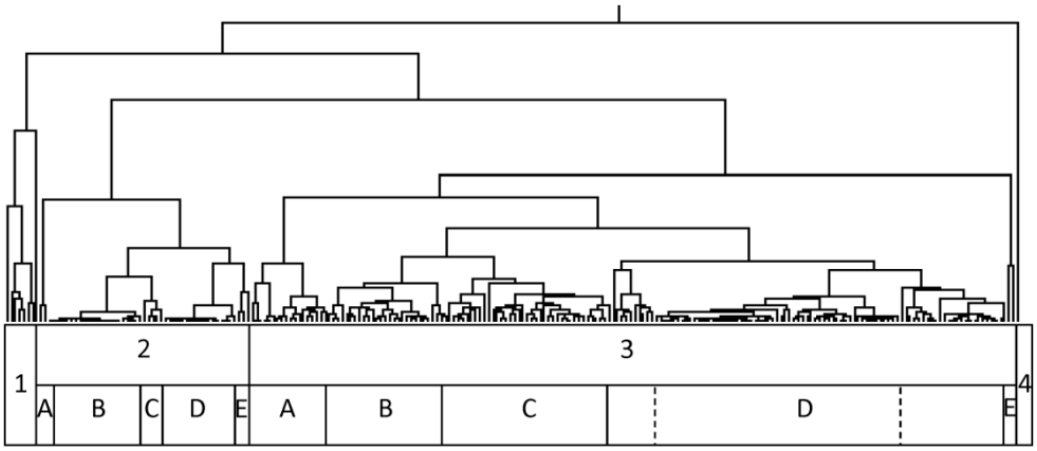


Figure 11: The portion of the dendrogram corresponding to the Aliphatic Cluster of samples. The labels were added to aid with discussion of each cluster. Subcluster D of Cluster 3 was subsequently divided with the dashed lines to indicate the left, middle, and right portions.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 2: A summary of compounds found within each subcluster of the aliphatic cluster. Carbon ranges are given in brackets.

Aliphatic Cluster		
Subcluster	Compounds	ASTM Class
1	[C_5-C_9] normal alkanes, cycloalkanes, and isoalkanes	
2A	[C_6-C_8] isoalkanes	ISO
2B	[C_8-C_{13}] isoalkanes	
2C	[C_9-C_{21}] isoalkanes	
2D	[C_{10}-C_{20}] normal alkanes	NA
2E	[C_7-C_{15}] isoalkanes; predominantly branched trimethylhexanes	ISO
3A	[C_6-C_{10}] normal alkanes, cycloalkanes, isoalkanes, and aromatics	PD
3B	[C_7-C_{17}] normal alkanes, cycloalkanes, isoalkanes, aromatics (alkanes:aromatics 2:1), and polynuclear aromatics	
3C	[C_7-C_{26}] normal alkanes, cycloalkanes, isoalkanes, aromatics (alkanes:aromatics 2:1), and polynuclear aromatics	
3D	Left: [C_8-C_{16}] normal alkanes, cycloalkanes, isoalkanes, aromatics (alkanes:aromatics 4:1), and polynuclear aromatics	
	Middle: [C_8-C_{12}] normal alkanes, cycloalkanes, isoalkanes, aromatics (alkanes:aromatics 4:1), and polynuclear aromatics	
	Right: [C_8-C_{17}] cycloalkanes and isoalkanes	NP
3E	[C_7-C_{12}] normal alkanes, cycloalkanes, and aromatics	PD
4	[C_{12}-C_{22}] polynuclear aromatics with [C_{15}-C_{22}] normal alkanes (99% weathered gasoline sample)	

The portion of the dendrogram in Figure 10 corresponding to the Aliphatic Cluster is shown in Figure 11 and can be further divided into four clusters. Table 2 summarizes the compounds present in each cluster and the carbon ranges shown encompass the lowest and highest carbon number among all samples in the clusters. In Table 2, samples in Cluster 1 contains normal alkanes, cycloalkanes, and isoalkanes within the C_5-C_9 carbon range. Cluster 2 can be split into five subclusters where A-C are isoalkanes with light, medium, and heavy carbon ranges, respectively. Subcluster D contains only normal alkanes. Subcluster E also contains isoalkanes; however, these are predominantly trimethylhexanes which have a molecular weight of 128. Therefore, the spectra contain *m*/*z* 128 which is indicative of polynuclear aromatics even though these liquids do not contain polynuclear aromatic compounds. Liquids in Cluster 2 correspond to the ISO and NA classes of the ASTM classification scheme. Cluster 3 contains

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

five subclusters where A is comprised of normal alkanes, cycloalkanes, isoalkanes, and aromatics in the carbon range of C_6-C_{10}. Subclusters B and C include the same compounds as A, where the alkane-to-aromatic ratio is 2:1 and polynuclear aromatics are also present. The ratios were calculated using the summed intensities for ions based on Table 2 of ASTM E1618-11 for alkanes (m/z 57, 71, 85, and 99) and aromatics (m/z 91, 105, 119, and 134). The left and middle portions of subcluster D are separated based on carbon ranges, but both contain the same compounds as B and C. An additional distinction for samples in these portions of subcluster D is the alkane-to-aromatic ratio which is at least 4:1. The right portion of subcluster D contains only cycloalkanes and isoalkanes and corresponds to the NP ASTM class. Subclusters A-D, with the exception of the right portion of D, corresponds to the PD ASTM class. Subcluster E contains cycloalkanes, normal alkanes, and aromatics in the carbon range of C_7-C_{12}. The sample labeled as Cluster 4, a 99% weathered gasoline, contains methyl naphthalenes with heavy normal alkanes.

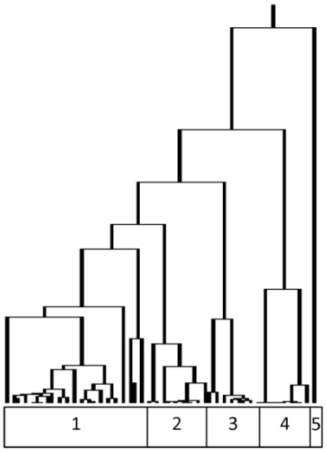


Figure 12: The portion of the dendrogram corresponding to the Aromatic Cluster of samples. The labels were added to aid with discussion of each cluster.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 3: A summary of compounds found within each subcluster of the aromatic cluster. Carbon ranges are given in brackets.

Aromatic Cluster		
Subcluster	Compounds	ASTM Class
1	[C_5-C_9] normal alkanes, isoalkanes, cyclopentanes, and cyclohexanes	GAS
	[C_7-C_{13}] toluene, C_2-, C_3-, and C_4-alkylbenzenes, and polynuclear aromatics	
2	[C_8-C_{12}] C_2- and C_3-alkylbenzenes	AR
3	[C_8-C_{14}] C_4-alkylbenzenes and polynuclear aromatics	
4	Left: [C_8-C_9] C_2-alkylbenzenes	
	Right: [C_7-C_{11}] toluene, C_2-alkylbenzenes may be present	
5	Aromatic with permethrin insecticide	

Figure 12 shows the portion of the dendrogram corresponding to the Aromatic Cluster. This can be further divided into five clusters. Table 3 provides a summary of the types of compounds present in each cluster. The carbon ranges shown in Table 3 encompass the lowest and highest carbon number among all samples in the clusters. In Table 3, samples in Cluster 1 contain aliphatic compounds with carbon ranges of C_6-C_9 (normal alkanes, isoalkanes, cyclopentanes, and cyclohexanes) and aromatic compounds in the range of C_7-C_{13} (toluene, C_2-, C_3-, and C_4-alkylbenzenes, and methyl naphthalenes). All samples corresponding to the GAS ASTM class were placed in Cluster 1, with the exception of three highly weathered gasolines that were placed in other clusters. Cluster 2 has a carbon range of C_8-C_{12} and includes C_2- and C_3-alkylbenzenes. Samples in Cluster 3 have a carbon range of C_8-C_{14} and contain C_4-alkylbenzenes and methyl naphthalenes. The left portion of Cluster 4 contains only C_2-alkylbenzenes, with a carbon range of C_8-C_9. All samples in the right portion of Cluster 4 have a carbon range of C_7-C_{11} and contain toluene; C_2-alkylbenzenes are also present in some samples. The sample labeled as Cluster 5 consists of aromatic compounds with permethrin (an insecticide). Clusters 2-5 correspond to the AR ASTM class and also include two highly weathered gasolines.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Hierarchical cluster analysis, an unsupervised learning method, was applied to IL TIS data from all ASTM classes other than the OXY class and MISC category. This technique resulted in clusters based on the chemical composition of the liquids. The clusters were found to be consistent with the classes outlined in ASTM E1618-11, where visual pattern recognition of the TIC is predominantly utilized. The TIS is conducive for development of multivariate statistical methods, automated search routines, and allows for interlaboratory comparisons that are not easily implemented based on the TIC.

Self-Organizing Feature Maps

The trained SOFM is shown in Figure 13a. The neurons are color-coded according to the ASTM designation(s) of the samples for which each neuron is the winning neuron (see color legend beside Figure 13b). For example, the neuron in the bottom left corner is shaded dark green to reflect it is the winning neuron for nine ignitable liquid samples all designated as MAR. This neuron will be referred to as one associated with the MAR (or more generally the AR) class. If samples of different ASTM class designations have the same winning neuron, the neuron is colored to reflect the fraction of samples for each class. For instance, the neuron in the upper left corner is colored to reflect that half of the two samples assigned to this neuron are designated as MISO and the other half as HISO. This neuron is associated with both the MISO and HISO (or more generally the ISO) classes. A neuron shaded with the same gray as the background of the figure is not associated with a training sample, and therefore no ASTM class is assigned to the neuron. This is referred to as an empty neuron.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

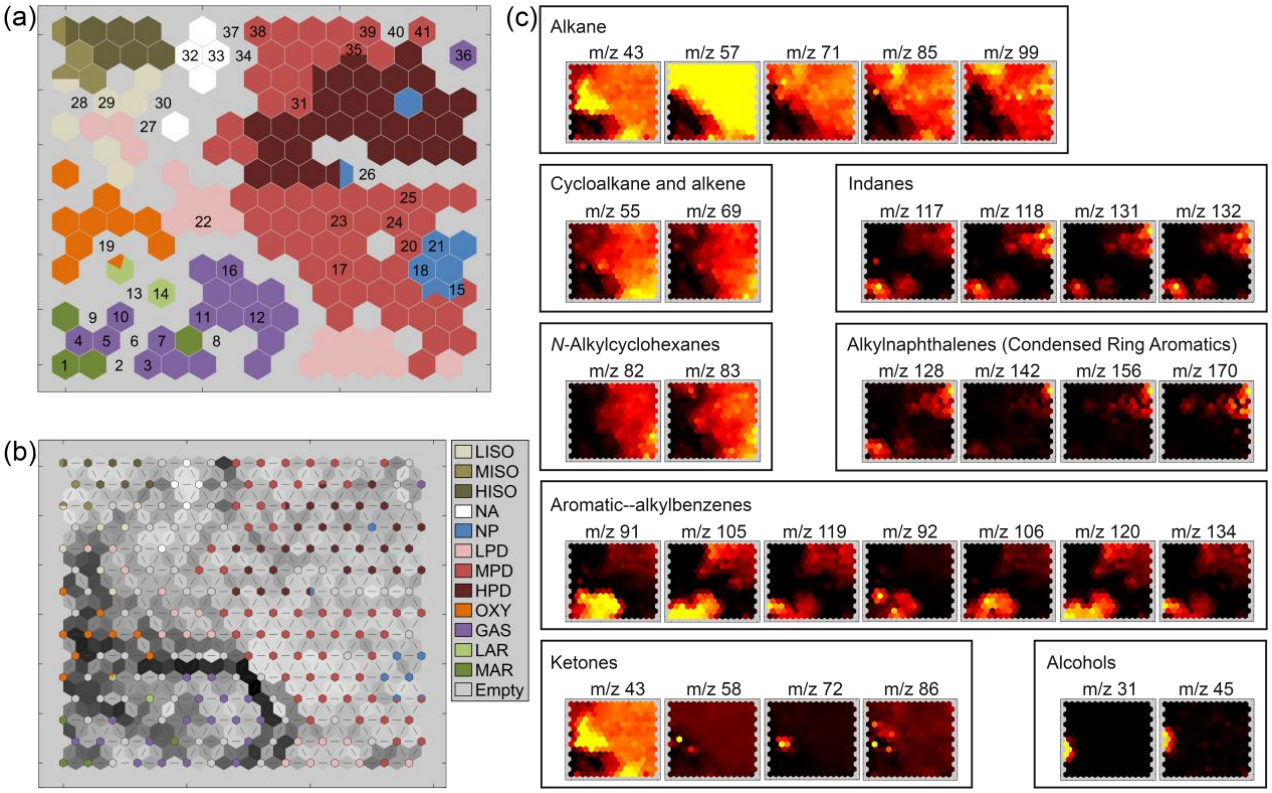


Figure 13: (a) The trained self-organizing feature map, where the neurons are colored according to their ASTM class associations. See the color-legend in Figure 1b. (b) Color-legend and U-matrix demonstrating relative distances between neurons in the trained map. The darker shaded areas between neurons indicate larger distances and the lighter shaded areas indicate shorter distances. The neurons are colored according to their ASTM class associations. (c) Component planes for ions in the extracted ion spectra. The component planes are grouped according to their common ion type. The lighter shaded neurons indicate high weight components while the darker shaded neurons indicate low weight components.

Figure 13a shows general grouping of neurons corresponding to ASTM class designations (GAS and AR, OXY, ISO, NA, NP and PD). However, when subclasses based on carbon range are considered, these subclasses are observed to be separated into different groups of neurons (*i.e.* two groups of MPD neurons). The cluster structure of the high-dimensional input data may be visualized in a two-dimensional unified distance matrix (or U-matrix) which shows the normalized distances between the weight vector of each neuron and neighboring neurons (80). The U-matrix for the calculated SOFM is shown in Figure 13b, where neurons are small hexagons with the same color-coding as Figure 13a. The lines illustrate the connections

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

between neurons, and the shaded regions surrounding each line indicate the normalized distance between the neurons. The darker regions correspond to the larger distances; whereas, lighter areas represent shorter distances where the neuron weight vectors are more similar. The darker shaded regions in Figure 13b indicate cluster borders which separate the cluster of neurons associated with AR, GAS, and OXY classes from those of ISO, NA, NP, and PD. The dark shaded areas around the OXY neurons demonstrate that not all of the OXY samples are spectrally similar to one another. A separation is also observed for the neurons associated with ISO and NA classes from those of PD and NP.

The groupings of neurons by ASTM class designations are further explored through visual analysis of the component planes, where each plane represents a relative comparison of an individual weight vector component for each neuron in the SOFM (27). The component planes for each of the input variables (Table 4) are shown in Figure 13c and are grouped according to the type of compound which typically produces the ion in EI mass spectra. The neurons in the component planes that are shaded with the brighter colors indicate a high weight while those shaded a darker color indicate a low weight. Relationships between classes of samples and their variables (ions) may be found by comparing the colors of the neurons in the component planes. Component planes with similar relative weights indicate the variables are correlated (81), as observed for each compound type in Figure 13c.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 4: Compound Type with Associated Ions Commonly Observed in Ignitable Liquids

Compound Type	m/z
Alkane	43, 57, 71, 85, 99
Cycloalkane and alkene	55, 69
n-Alkylcyclohexanes	82, 83
Aromatic—alkylbenzenes	91, 105, 119, 92, 106, 120, 134
Indanes	117, 118, 131, 132
Alkylnaphthalenes (Condensed Ring Aromatics)	128, 142, 156, 170
Ketones	43, 58, 72, 86
Alcohols	31, 45

The EIS corresponding to GAS and AR samples are clustered in the lower left corner of the SOFM, except for the triplicate analyses of 99% weathered GAS samples. The separation of this cluster from the other ASTM class associated neurons is reflected throughout the alkane and aromatic component planes. It is clear that GAS and AR associated neurons have higher aromatic ion weights and lower alkane ion weights compared to the SOFM neurons associated with the other ASTM classes. According to ASTM E1618-11, GAS samples in the United States generally have a higher concentration of aromatic compounds compared to alkanes, and samples in the AR class are almost exclusively comprised of aromatic compounds (6). Twenty-four of the 38 GAS samples included in the training data set were weathered from a range of 25% to 99%. A GAS sample was analyzed at 0% (unweathered), 50%, 75%, and 90% weathered. The EIS of the unweathered GAS sample and the EIS of the 50% weathered sample are assigned to the same neuron, demonstrating that their EIS are similar. The EIS for the 75% and 90% weathered GAS samples are each assigned to two separate neurons. The neuron associated with GAS in the upper right-hand corner of the SOFM is associated with a 99% weathered gasoline, and shows high component weights for alkanes (*m/z* 57 and 99), indanes (*m/z* 118), and each of the condensed ring aromatics (*m/z* 128, *m/z* 142, *m/z* 156, and *m/z* 170). Twenty-four AR

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

samples (16 MAR and 8 LAR) are separated in the SOFM based on their carbon range, and the 16 MAR samples are assigned to different neurons based on the presence or absence of naphthalenes.

Located above the GAS and AR associated neurons in the SOFM are the OXY associated neurons. The 16 samples designated as OXY in the training data set are associated with nine neurons, and one of those samples containing a significant toluene peak is assigned to a neuron primarily associated with LAR. High weights in the component planes for alcohols and ketones are observed for those OXY associated neurons. One of these component planes is *m/z* 43 which also corresponds to alkane compounds, as shown in Figure 13c. Samples containing predominantly alcohols were assigned to different neurons than those containing ketones; however, the neurons do not appear to be clustered in the SOFM according to these trends. The 35 ISO training samples (7 HISO, 20 MISO, and 8 LISO) are assigned to neurons in the upper left corner of the SOFM. ISO samples contain almost exclusively branched chain aliphatic compounds (6), which is reflected in the high weights of the alkane component planes.

Seventeen NA samples, comprised exclusively of *n*-alkanes (6), are assigned to five neurons with high weights in the alkane component planes. The NA neuron located slightly apart from the others is associated with NA samples having slightly lower EIS intensities for *m/z* 71 and *m/z* 85 when compared to the other NA samples.

Petroleum distillate samples contain alkanes, cycloalkanes, aromatics, and may contain condensed ring aromatic compounds, where alkanes are the most abundant (6). The SOFM neurons associated with the PD class have higher component plane weights for the ions associated with alkanes, cycloalkanes, and *n*-cyclohexanes compared to neurons associated with the ISO, NA, and NP classes. The component planes also indicate the presence of aromatics in

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

some of the PD associated neurons. A group of neurons is observed for the 48 samples designated as HPD, which divides the MPD associated neurons into two groups. Of the 94 MPD samples in the training set, those with a higher concentration of aromatic compounds (traditional distillates (6)) are assigned to a group of neurons located above the HPD group, while those with a lower concentration of aromatics (de-aromatized (6)) are located below the HPD associated neurons. The traditional distillate and de-aromatized MPD groups can be distinguished from one another by comparing the component planes for the indanes, condensed ring aromatics, and aromatic compound types. The 25 samples designated as LPD are assigned to two groups of neurons in the SOFM. The cycloalkane and alkene component planes reflect the differences between the two groups as heavier weights are observed for the group located near the bottom of the SOFM.

The final ASTM class is NP, comprised of cyclic and branched alkanes (6), which is demonstrated in the weight components of the *n*-alkylcyclohexanes and cycloalkane compound types. Of the 16 designated NP samples in the data set, one (SRN 395) is assigned to a neuron surrounded by HPD associated neurons, and its EIS is observed to have a higher intensity of aromatic associated ions than the remaining NP samples. This liquid has a consensus classification of NP in the ILRC. The liquid contains a small quantity of 1,2,4-trimethylbenzene, which is allowed under the ASTM E1618 guidelines. Nine of the NP samples are assigned to four neurons which are clustered together, while the remaining samples are assigned to neurons also associated with HPD or MPD classes. The EIS data for the samples associated with these neurons are similar.

Auto-scaling the EIS before SOFM calculations did not improve the results described in the previous paragraphs. Clustering results from the SOFM which rely on a subset of ions in a

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

total ion spectrum were similar to results obtained by hierarchical cluster analysis (HCA) of ignitable liquids using total ion spectra. Hierarchical cluster analysis of ignitable liquid TIS resulted in clusters based on the chemical composition of the liquids (8). The SOFM study used fewer ions, attempted to cluster ignitable liquids with multiple ASTM class characteristics, and provided an enhanced representation of the class relationships compared to the HCA study.

The EIS for 116 fire debris samples were projected onto the SOFM. The winning neurons are numbered in Figure 13a, and the results are summarized in Table 5. Three sample neuron association types are observed: "match", "empty-adjacent", and "mis-match". When the ASTM class designation for the fire debris sample is the same as the ASTM class associated with the neuron where the fire debris sample was projected, the projection is identified as a "match". This indicates that the variables (ions) for the fire debris sample are similar to those of the ignitable liquid sample(s) assigned to that neuron during training. For example, two of the fire debris samples designated as LISO were projected to the neuron labeled as 29, which was associated with the LISO class during training. When the winning neuron is empty and adjacent to a neuron with the same class association as the fire debris sample and the two neurons have similar component weights, this is identified as "empty-adjacent". For example, 21 fire debris GAS samples were assigned to empty neurons labeled 2, 6, 8, 9, and 13, which are adjacent to GAS associated neurons. Neurons labeled 2, 9 and 13 are also adjacent to aromatic-associated neurons, which is expected given the aromatic content of gasoline. Analysis of the component planes in Figure 13c show that these "empty-adjacent" neurons have many similar component weights as those associated with GAS samples; however, the assignment of the fire debris sample to an empty neuron indicates a variation in the EIS data compared to those gasoline samples in the training data set. Seven of the 116 fire debris samples are projected to neurons

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

with a different class association, and these are identified as a "mis-match". For instance, two samples designated as HPD are projected to the neuron labeled as 36, which was associated with the 99% weathered GAS. This neuron does share many similar component weights as other HPD associated neurons.

Table 5: Comparison of Sample Designations and Assigned Neuron Associations for Fire Debris Samples Projected into the SOFM

Fire Debris Sample Designation[a]	"Match"[b]		"Empty-Adjacent"[c]		"Mis-Match"[d]	
	No. Samples	Neuron No.[e]	No. Samples	Neuron No.[e]	No. Samples	Neuron No.[e]
LISO	2	29	2	27, 30	-	-
MISO	7	28	-	-	-	-
NA	2	32, 33	5	34, 40	-	-
NP	7	18, 21, 15*	-	-	1	39
LPD	1	22	-	-	1	39
MPD	16	17, 20, 23, 24, 25, 31, 38, 41, 35*	1	26	3	18, 39
HPD	-	-	-	-	2	36
GAS	36	3, 4, 5, 7, 10, 11, 12, 16	21	2, 6, 8, 9, 13	-	-
LAR	3	14	-	-	-	-
MAR	6	1	-	-	-	-

[a] See text for abbreviation definitions.
[b] The ASTM class designation for the fire debris sample(s) is the same as the class association of the winning neuron, indicated by the neuron number.
[c] The winning neuron was considered to be empty but was adjacent to a neuron with the same class association as the ASTM class designation of the fire debris sample(s).
[d] The ASTM class designation for the fire debris sample(s) is not the same as the class association of the winning neuron, indicated by the neuron number
[e] The neuron number corresponds to the neuron label in Figure 1a.
*Neuron also associated with multiple ASTM classes, see first paragraph in Section 3.1.

Finally, the EIS for select ignitable liquid samples designated as OXY or MISC are projected onto the SOFM. The 33 samples designated as MISC, have spectral characteristics of multiple ASTM classes (*i.e.* HPD+AR, LPD+MPD, or NP+NA). Of these samples, 22 are projected to neurons that are associated with one of the ASTM classes or to empty neurons that

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

are adjacent to a neuron associated with one of the classes. The winning neurons of six samples are empty and are not adjacent to neurons associated with one of the classes; while, five samples are projected into neurons associated with other classes than those of the liquid. Seven samples designated as OXY, which contain mixtures of non-oxygenated compounds with ASTM class characteristics (*i.e.* LPD+acetone), are projected onto the SOFM. Projection of these samples onto the SOFM results in four of the samples being assigned to a neuron associated with the non-oxygenated class, while one sample (LPD+acetone) is assigned to an empty neuron adjacent to both OXY and LPD associated neurons. Two of the samples are assigned to empty neurons that are not adjacent to neurons associated with either the OXY or the non-oxygenated classes.

Discriminant Analysis

LDA and QDA

Retaining a minimum of 95% of the variance within the data were found to yield the highest correct classification percentages. The number of principal components retained for model development was based on this value.

For each step of the classification scheme, the covariance matrices were determined to be statistically different, indicating QDA is the more appropriate analysis method. Results for the less appropriate LDA method will be compared to those using QDA to understand the influence of the model (LDA or QDA) on the correct classification rates. The same multi-step process and data sets were used for testing both the LDA and QDA models. The results discussed below are for the data set that was normalized so that each sample's intensities summed to one.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The true positive rate and false positive rate are shown in Table 6 for the first step of the multi-step scheme (classification into the IL and SUB classes). The values in parentheses represent the metrics for the data set with 20% substrate contribution. The level of substrate contribution is uncontrolled in the fire debris samples containing an ignitable liquid; therefore, the error rates under the "Fire Debris" heading reflect errors based on using models constructed with data having an upper limit of 20% substrate contribution. The false positive rate should ideally be kept low and this important rate may dictate the choice of the best model. The QDA model produced the lowest false positive rate in cross-validation data sets, Table 6. Similarly, QDA gave the lowest false positive rate for the fire debris samples as well; however, the false positive rate increased when the fire debris sample designations were based on the analyst. The cross-validation results demonstrate that the QDA model can maintain low false positive rate for data sets that resemble the data used to create the model. The fire debris results, with designations based on the pour, demonstrate a lower false positive rate than when the designations were based on the analyst.

The different behavior for the models applied to the cross-validation and fire debris data can be understood in terms of the Mahalanobis distance. When the substrate contribution of the training data is increased, the model adjusts to classify samples with high levels of substrate contribution into the IL class. Increased classification into the IL class for a given sample occurs because the multivariate probability ellipsoid for the IL class increases in size and results in a reduced Mahalanobis distance between a test sample and the IL class, while the size of the probability ellipsoid for the SUB class is unchanged. Models based on data sets containing smaller substrate contribution will result in an IL class with a smaller probability ellipsoid. Test samples falling between the IL and SUB probability ellipsoids will be classified based on the

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

smallest Mahalanobis distance to the IL or SUB class. For this reason, a model with some substrate contribution may show improved classification performance; however, if the contribution is too large, model classification performance will decrease.

Table 6: Classifier Performance Metrics for LDA and QDA

	Cross-Validation			Fire Debris - Pour			Fire Debris - Analyst	
Classifier	LDA	QDA		LDA	QDA		LDA	QDA
True Positive Rate	95.5 (94.0)	81.3 (86.5)		81.0 (81.6)	70.9 (81.0)		96.7 (94.3)	84.6 (94.3)
False Positive Rate	24.5 (12.1)	9.9 (6.0)		15.6 (17.8)	8.9 (17.8)		20.0 (26.3)	15.0 (25.0)
True Negative Rate	75.5 (87.9)	90.1 (94.0)		84.4 (82.2)	91.1 (82.2)		80.0 (73.7)	85.0 (75.0)
False Negative Rate	4.5 (6.0)	18.7 (13.5)		19.0 (18.4)	29.1 (19.0)		3.3 (5.7)	15.4 (5.7)
Precision	95.2 (89.7)	97.7 (94.2)		94.8 (94.2)	96.6 (94.1)		88.1 (84.7)	89.7 (85.3)
Accuracy	92.2 (91.1)	82.7 (90.0)		81.8 (81.8)	75.4 (81.3)		90.2 (86.2)	84.7 (86.7)

0% Substrate Contribution

The LDA and QDA results are shown in confusion matrices for the cumulative cross-validation test set, Figure 14, and the fire debris samples with designations based on the proximity to the pour and the analyst, Figure 15 and Figure 16, respectively. The QDA values are shown in parentheses. The first column in Figure 14 represents the designated class and the column headings represent the class assigned by the model. The entries in the table give the number of cross validation samples assigned to each class. For example, out of 9,200 ignitable liquid cross-validation samples (i.e. 460 IL * 20%/iteration*100 iterations=9200), 8,784 were correctly assigned to IL, while 416 cross-validation samples were incorrectly assigned to SUB using LDA. This represents the true positive rate of 95.5% (LDA) in Table 6 for the 0% substrate contribution cross-validation. The fire debris results in Figure 15 reflect a drop in the true positive rate to 81.0% (LDA). This drop in true positive rate may be attributed to the loss of

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

ignitable liquid residue during the burning process. The LDA false positive rate drops from 24.5% (cross-validation) to 15.6% (fire debris), while the QDA false positive rate remains at 9-10% for both cross-validation and fire debris samples.

	IL	SUB	Total	%Correct
IL	8,784 (7,480)	416 (1,720)	9,200	95.5 (81.3)
SUB	441 (178)	1,359 (1,622)	1,800	75.5 (90.1)
Total	9,225 (7,658)	1,775 (3,342)	11,000	92.2 (82.7)

	ALI	ARG	Total	%Correct
ALI	4,700 (4,664)	0 (36)	4,700	100.0 (99.2)
ARG	62 (53)	1,138 (1,147)	1,200	94.8 (95.6)
Total	4,762 (4,717)	1,138 (1,183)	5,900	98.9 (98.5)

	AR	GAS	Total	%Correct
AR	393 (495)	107 (5)	500	78.6 (99.0)
GAS	104 (17)	696 (783)	800	87.0 (97.9)
Total	497 (512)	803 (788)	1,300	83.8 (98.3)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	942 (956)	58 (44)	1,000	94.2 (95.6)
PD/NP	8 (79)	3,692 (3,621)	3,700	99.8 (97.9)
Total	950 (1,035)	3,750 (3,665)	4,700	98.6 (97.4)

	ISO	NA	Total	%Correct
ISO	572 (663)	128 (37)	700	81.7 (94.7)
NA	13 (24)	287 (276)	300	95.7 (92.0)
Total	585 (687)	415 (313)	1,000	85.9 (93.9)

	PD	NP	Total	%Correct
PD	2,520 (3,051)	780 (249)	3,300	76.4 (92.5)
NP	3 (24)	297 (276)	300	99.0 (92.0)
Total	2,523 (3,075)	1,077 (525)	3,600	78.3 (92.4)

Figure 14: Cumulative results for cross-validation using the data set with 0% substrate contribution. The QDA results are shown in parentheses. The total number of samples used for LDA and QDA was the same.

In Figure 15, confusion matrices are given for LDA and QDA classification of the fire debris samples, with designations based on the proximity to the pour. The QDA values are shown in parentheses. These results were obtained using the 0% SUB contribution data set as the model. As with the previous confusion matrices, the row headings indicate the designated class, and the column headings give the class assigned by the model.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Total	%Correct
IL	128 (112)	30 (46)	158	81.0 (70.9)
SUB	7 (4)	38 (41)	45	84.4 (91.1)
Total	135 (116)	68 (87)	203	81.8 (75.4)

	ALI	ARG	Total	%Correct
ALI	55 (39)	11 (27)	66	83.3 (59.1)
ARG	5 (0)	77 (82)	82	93.9 (100.0)
Total	60 (39)	88 (109)	148	89.2 (81.8)

	AR	GAS	Total	%Correct
AR	8 (9)	1 (0)	9	88.9 (100.0)
GAS	16 (7)	57 (66)	73	78.1 (90.4)
Total	24 (16)	58 (66)	82	79.3 (91.5)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	13 (8)	6 (11)	19	68.4 (42.1)
PD/NP	0 (0)	47 (47)	47	100.0 (100.0)
Total	13 (8)	53 (58)	66	90.9 (83.3)

	ISO	NA	Total	%Correct
ISO	9 (12)	3 (0)	12	75.0 (100.0)
NA	5 (7)	2 (0)	7	28.6 (0.0)
Total	14 (19)	5 (0)	19	57.9 (63.2)

	PD	NP	Total	%Correct
PD	33 (34)	3 (2)	36	91.7 (94.4)
NP	4 (6)	7 (5)	11	63.6 (45.5)
Total	37 (40)	10 (7)	47	85.1 (83.0)

Figure 15: Confusion matrices for fire debris samples based on models using 0% substrate contribution. Designations for the fire debris samples are based on the proximity to the pour.

Figure 16 shows the confusion matrices for the LDA and QDA classifications of the fire debris samples with designations based on the analyst. The procedure for assigning designations based on the analyst was previously discussed. The QDA results are given in parentheses. The correct classification rates obtained for these samples are very similar to those obtained for the samples with designations based on the proximity to the pour. The first two steps of the classification scheme reflect a higher correct classification rate using designations based on the analyst. As expected, this designation method reduces the number of false negatives; however, the false positive rate increased. This indicates that the classification model associates components in these samples with ILs even though the analyst did not identify an IL profile in the sample data.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Total	%Correct
IL	119 (104)	4 (19)	123	96.7 (84.6)
SUB	16 (12)	64 (68)	80	80.0 (85.0)
Total	135 (116)	68 (87)	203	90.1 (84.7)

	ALI	ARG	Total	%Correct
ALI	49 (39)	1 (11)	50	98.0 (78.0)
ARG	0 (0)	66 (66)	66	100.0 (100.0)
Total	49 (39)	67 (77)	116	99.1 (90.5)

	AR	GAS	Total	%Correct
AR	8 (9)	1 (0)	9	88.9 (100.0)
GAS	16 (7)	41 (50)	57	71.9 (87.7)
Total	24 (16)	42 (50)	66	74.2 (89.4)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	13 (8)	5 (10)	18	72.2 (44.4)
PD/NP	0 (0)	32 (32)	32	100.0 (100.0)
Total	13 (8)	37 (42)	50	90.0 (80.0)

	ISO	NA	Total	%Correct
ISO	8 (11)	3 (0)	11	72.7 (100.0)
NA	5 (7)	2 (0)	7	28.6 (0.0)
Total	13 (18)	5 (0)	18	55.6 (61.1)

	PD	NP	Total	%Correct
PD	21 (22)	3 (2)	24	87.5 (91.7)
NP	3 (3)	7 (5)	8	87.5 (62.5)
Total	22 (40)	10 (7)	32	87.5 (84.4)

Figure 16: Confusion matrices for fire debris samples based on models using 0% substrate contribution. Designations for the fire debris samples are based on the analyst's review of the post-burn data.

When LDA was applied to the cross-validation test set for IL/SUB discrimination, there was a total correct classification rate of 92.2% for a total of 11,000 samples. QDA was applied to the same size data set and a correct classification rate of 82.7% was observed. The cross-validation calculation was repeated 100 times, as described earlier, and each of the 100 iterations in the cross-validation utilized a different model data set and consequently resulted in a slightly different set of scores and latent variables from PCA. Retention of 95% of the variance in the model data required approximately 10 principal components. The first three principal components contained approximately 75% of the variance. PCA maximizes the fraction of the variance retained in each latent variable and the resulting scores may not necessarily provide a graphical visualization of class separation. The LDA results from the confusion matrices in Figure 14 and Figure 15 can be visualized by projection onto the single canonical variate

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

obtained from Fisher discriminant analysis of the data from the two classes. In Figure 17, each step of the classification scheme is represented graphically. The Gaussian curves in each figure represent the distributions of the projected training data from the two classes, while the symbols below the graph shows the distribution of the projected fire debris data, with designations based on proximity to the pour. The filled symbols correspond to the solid curve and the open symbols correspond to the dashed curve. The vertical lines in the fire debris distribution plots correspond to the crossing point for the two Gaussian curves. Points on either side of the vertical line are placed in the class of the overlying curve. For example, in Figure 17a, the solid curve and circle symbols correspond to the IL class and the dashed curve and triangle symbols correspond to the SUB class.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

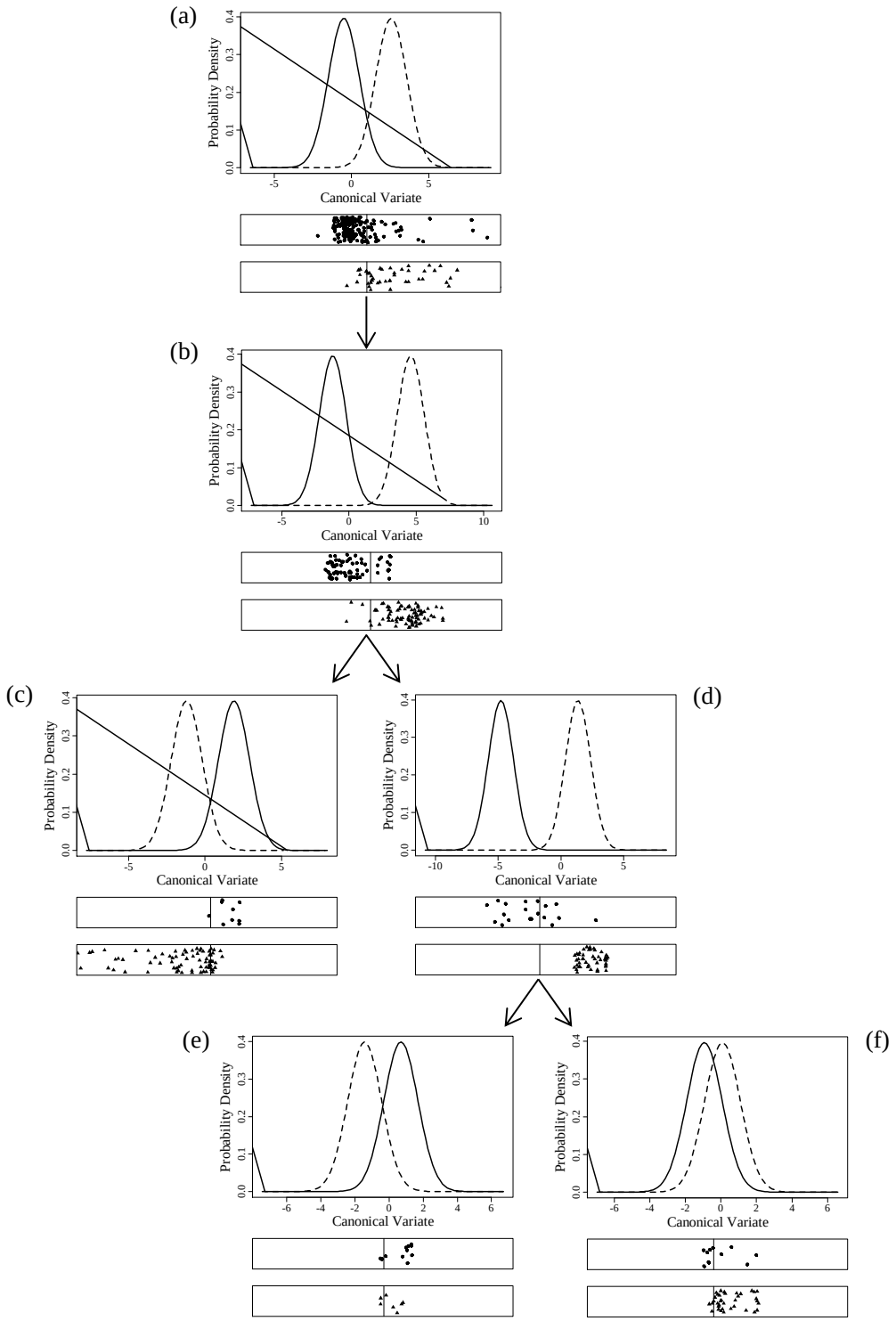


Figure 17: The Gaussian curves in a – f represent the distributions of the projected LDA training data from the two classes for each of the respective tables in Figure 2, while the symbols below the graph show the distributions of the projected fire debris data from the tables in Figure 3. Thecircle symbols correspond to the solid curve and the triangle symbols correspond to the dashed curve. The solid curve and filled symbols, and the dashed curve and open symbols correspond to (a) IL and SUB, (b) ALI and ARG, (c) AR and GAS, (d) ISO/NA and PD/NP, (e) ISO and NA, (f) NP and PD. See text for abbreviation definitions.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

When the models were tested on the fire debris samples, with designations based on the proximity to the pour, LDA gave a correct classification rate of 81.8%, 30 ignitable liquids misclassified as substrates, and seven substrate samples misclassified as ignitable liquids. The misclassification of IL as SUB may be attributed to total evaporation of the IL and demonstrates the need to know the ground truth in evaluating the accuracy of the method. Using QDA resulted in a correct classification rate of 75.4%, 46 ignitable liquids misclassified as substrates, and four substrates misclassified as ignitable liquids. When the designations for the fire debris samples were based on the analyst's review of the post-burn data, a correct classification rate of 90.1% was observed. The increase in correct classification rate is attributed to the analyst's conservative approach, requiring a stronger IL residue to determine a sample positive for IL. The stronger IL contribution results in a higher percent correct classification rate.

The next step was to classify an ignitable liquid sample as either an ARG or ALI sample. Correct classification rates of 98.9% for the cross-validation test set and 89.2% for the fire debris samples, with designations based on proximity to the pour, were obtained using LDA. When the designations were based on the analyst's review of the data, a correct classification rate of 99.1% was achieved. Using QDA, 98.5% of samples selected for the cross-validation test set classified correctly. The fire debris samples resulted in correct classification rates of 81.8% and 90.5% with designations based on the proximity to the pour and the analyst's review of the post-burn data, respectively.

The ARG samples were further classified into the ASTM E1618-10 AR or GAS classes. Using LDA, 83.8% of samples selected for the cross-validation test set and 79.3% of fire debris samples, with designations based on the pour, correctly classified. Using the designations based

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

on the analyst, a correct classification rate of 74.2% was obtained. Correct classification rates of 98.3% for the cross-validation test set and 91.5% for the fire debris samples, with designations based on the proximity to the pour, were obtained with QDA. When the designations were based on the analyst, 89.4% of fire debris samples classified correctly. Using QDA, all nine AR fire debris samples, for both designations, classified correctly, while a number of GAS samples misclassified as aromatics. This may be attributed to the heat of the fire causing evaporation of the lighter components in the gasoline samples. The post-burn profile will change significantly due to the remaining heavier aromatic compounds; therefore, caution should be used when trying to classify a post-burn ARG sample into its respective ASTM classification.

ALI samples were then classified as belonging to the combined ISO/ NA class or the combined PD/NP class. Cross-validation for this step had high correct classification rates of 98.6% using LDA and 97.4% using QDA when predicting classifications of 4,700 samples. The fire debris samples, with designations based on the proximity to the pour, had correct classification rates of 90.9% for LDA and 83.3% for QDA. Using the designations based on the analyst, the fire debris samples had correct classification rates of 90.0% and 80.0% for LDA and QDA, respectively.

Next, samples belonging to the ISO/NA or PD/NP combined classes were separated into their respective ASTM E1618-10 classes. The classification of the cross-validation test set into ISO or NA resulted in 85.9% correct classification when using LDA and 93.9% correct when using QDA. Significant limitations were seen in separating fire debris samples into the ISO or NA ASTM classes. For fire debris samples with designations based on the proximity to the pour, LDA gave a correct classification rate of 57.9% while QDA was at 63.2%. Using the designations based on the analyst, correct classification rates of 55.6% and 66.1% were obtained

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

for LDA and QDA, respectively. Classifying cross-validation samples into PD or NP using LDA gave a correct classification rate of 78.3%, while QDA gave 92.4% correct. The fire debris samples, with designations based on the proximity to the pour, had correct classification rates of 85.1% and 83.0% for LDA and QDA, respectively. When the designations for the fire debris samples were based on the analyst, a correct classification rate of 87.5% was obtained for LDA, while for QDA it was 84.4%.

20% Substrate Contribution

Confusion matrices show the LDA and QDA results for the cumulative cross-validation test set, Figure 18, and the fire debris samples with designations based on the proximity to the pour and the analyst's review of the post-burn data in Figure 19 and Figure 20, respectively. Projection of the LDA results onto the canonical variate is not shown for the data set with an upper limit of 20% substrate contribution.

	IL	SUB	Total	%Correct
IL	95,149 (87,518)	6,051 (13,682)	101,200	94.0 (86.5)
SUB	10,822 (5,351)	78,918 (84,449)	89,800	87.9 (94.0)
Total	106,031 (92,869)	84,969 (98,131)	191,000	91.1 (90.0)

	ALI	ARG	Total	%Correct
ALI	51,700 (51,500)	0 (200)	51,700	100.0 (99.6)
ARG	642 (524)	12,958 (13,076)	13,600	95.3 (96.1)
Total	52,342 (52,024)	12,958 (13,276)	65,300	99.0 (98.9)

	AR	GAS	Total	%Correct
AR	4,525 (5,300)	775 (0)	5,300	85.4 (100.0)
GAS	590 (34)	7,810 (8,366)	8,400	93.0 (99.6)
Total	5,115 (5,334)	8,585 (8,366)	13,700	90.0 (99.8)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	10,777 (11,269)	623 (131)	11,400	94.5 (98.9)
PD/NP	0 (898)	40,300 (39,402)	40,300	100.0 (97.8)
Total	10,777 (12,167)	40,923 (39,533)	51,700	98.8 (98.0)

	ISO	NA	Total	%Correct
ISO	6,589 (7,649)	1,111 (51)	7,700	85.6 (99.3)
NA	25 (22)	3,675 (3,678)	3,700	99.3 (99.4)
Total	6,614 (7,671)	4,786 (3,729)	11,400	90.0 (99.4)

	PD	NP	Total	%Correct
PD	27,861 (33,011)	8,839 (3,689)	36,700	75.9 (89.9)
NP	39 (146)	3,461 (3,354)	3,500	98.9 (95.8)
Total	27,900 (33,157)	12,300 (7,043)	40,200	77.9 (90.5)

Figure 18: Cumulative results for cross-validation using the data set with an upper limit of 20% substrate contribution. The QDA results are shown in parentheses. The total number of samples used for LDA and QDA was the same.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Total	%Correct
IL	129 (128)	29 (30)	158	81.6 (81.0)
SUB	8 (8)	37 (37)	45	82.2 (82.2)
Total	137 (136)	66 (67)	203	81.8 (81.3)

	ALI	ARG	Total	%Correct
ALI	54 (44)	12 (22)	66	81.8 (66.7)
ARG	5 (1)	77 (81)	82	93.9 (98.8)
Total	59 (45)	89 (103)	148	88.5 (84.5)

	AR	GAS	Total	%Correct
AR	9 (9)	0 (0)	9	100.0 (100.0)
GAS	26 (9)	47 (64)	73	64.4 (87.7)
Total	35 (18)	47 (64)	82	68.3 (89.0)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	15 (14)	4 (5)	19	78.9 (73.7)
PD/NP	0 (0)	47 (47)	47	100.0 (100.0)
Total	15 (14)	51 (52)	66	93.9 (92.4)

	ISO	NA	Total	%Correct
ISO	8 (10)	4 (2)	12	66.7 (83.3)
NA	4 (5)	3 (2)	7	42.9 (28.6)
Total	12 (15)	7 (4)	19	57.9 (63.2)

	PD	NP	Total	%Correct
PD	30 (33)	6 (3)	36	83.3 (91.7)
NP	4 (5)	7 (6)	11	63.6 (54.5)
Total	34 (38)	13 (9)	47	78.7 (83.0)

Figure 19: Confusion matrices for fire debris samples based on models using an upper limit of 20% substrate contribution. Designations for the fire debris samples are based on the proximity to the pour.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Total	%Correct
IL	116 (116)	7 (7)	123	94.3 (94.3)
SUB	21 (20)	59 (60)	80	73.8 (75.0)
Total	137 (136)	66 (67)	203	86.2 (86.7)

	ALI	ARG	Total	%Correct
ALI	48 (41)	2 (9)	50	96.0 (82.0)
ARG	0 (0)	66 (66)	66	100.0 (100.0)
Total	48 (41)	68 (75)	116	98.3 (92.2)

	AR	GAS	Total	%Correct
AR	9 (9)	0 (0)	9	100.0 (100.0)
GAS	26 (9)	31 (48)	57	54.4 (84.2)
Total	35 (18)	31 (48)	66	60.6 (86.4)

	ISO/NA	PD/NP	Total	%Correct
ISO/NA	15 (14)	3 (4)	18	83.3 (77.8)
PD/NP	0 (0)	32 (32)	32	100.0 (100.0)
Total	15 (14)	35 (36)	50	94.0 (92.0)

	ISO	NA	Total	%Correct
ISO	7 (9)	4 (2)	11	63.6 (81.8)
NA	4 (5)	3 (2)	7	42.9 (28.6)
Total	11 (14)	7 (4)	18	55.6 (61.1)

	PD	NP	Total	%Correct
PD	18 (21)	6 (3)	24	75.0 (87.5)
NP	1 (2)	7 (6)	8	87.5 (75.0)
Total	19 (23)	13 (9)	32	78.1 (84.4)

Figure 20: Confusion matrices for fire debris samples based on models using an upper limit of 20% substrate contribution. Designations for the fire debris samples are based on the analyst's review of the post-burn data.

For the first step of the classification scheme, a total correct classification rate of approximately 91.1% for 191,000 samples was observed for the cross-validation test set when using LDA. A correct classification rate of 90.0% was observed for the same size data set when using QDA.

A correct classification rate of 81.8% was observed when testing the LDA model with fire debris samples, with designations based on proximity to the pour, 29 ignitable liquids misclassified as substrates, and eight substrate samples misclassified as ignitable liquids. Using the QDA model, a correct classification percentage of 81.3% was obtained, 30 ignitable liquids misclassified as substrates, and eight substrates misclassified as ignitable liquids. When the designations for the fire debris samples were based on the analyst's examination of the post-burn data, a correct classification rate of 86.2% was obtained with LDA. There were seven IL

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

samples that misclassified as SUB, and 21 SUB samples that misclassified as IL. Using QDA, a correct classification rate of 86.7% was achieved. There were seven IL samples that misclassified as SUB, and 20 SUB samples that misclassified as IL.

For classifying ignitable liquid samples as belonging to the ALI or ARG classes, 99.0% of the samples in the cross-validation test set were correctly classified, and 88.5% of the fire debris samples, with designations based on the pour, were correctly classified when using LDA. Correct classification rates of 98.9% of the samples selected for the cross-validation test set and 84.5% of the fire debris samples, with designations based on the pour, were observed when using QDA. When the designations were based on the analyst's review of the post-burn data, correct classification rates of 98.3% and 92.2% were obtained for LDA and QDA, respectively.

When the ARG samples were classified into their individual classes, LDA gave correct classification rates of 90.0% for the cross-validation test set and 68.3% for the fire debris samples, with designations based on the pour. The fire debris samples with designations based on the analyst achieved a correct classification rate of 60.6%. Using QDA gave correct classification rates of 99.8% for the cross-validation test set and 89.0% for the fire debris samples, with designations based on the pour. When the designations were based on the analyst, a correct classification rate of 86.4% was obtained. For both designations, all nine aromatic fire debris samples classified correctly using LDA and QDA, while a number of gasoline samples misclassified as aromatics. This can be attributed to reasons previously discussed for the model developed with 0% substrate contribution.

Next, ALI samples were classified as ISO/ NA or PD/NP. Cross-validation using both LDA and QDA for this step had high correct classification rates of approximately 98% when predicting classifications of 51,700 samples. The fire debris samples, with designations based on

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

the pour, had correct classification rates of 93.9% and 92.4% for LDA and QDA, respectively. Using the designations based on the analyst, correct classification rates of 94.0% and 92.0% were achieved for LDA and QDA, respectively.

The classification of the cross-validation test set into ISO or NA resulted in 90.0% correct classification when using LDA and 99.4% correct when using QDA. As was also observed for the model developed with 0% substrate contribution, separating fire debris samples designated as ISO/ NA showed significant limitations. For fire debris samples, with designations based on the pour, LDA gave a correct classification rate of 57.9% while QDA was at 63.2%. The fire debris samples with designations based on the analyst resulted in correct classification rates of 55.6% and 61.1% for LDA and QDA, respectively. Classifying cross-validation samples into PD or NP using LDA gave a correct classification rate of 77.9%, while QDA gave 90.5% correct. The fire debris samples, with designations based on the pour, had correct classification rates of 78.7% and 83.0% for LDA and QDA, respectively. When the designations were based on the analyst, correct classification rates of 78.1% and 84.4% were obtained for LDA and QDA, respectively.

SIMCA

Prior to cross-validation, PCA was performed on each class of the data set used in the given step of the classification scheme and outliers were identified and removed if they exceeded the 97.5 percentile of either the score distance or orthogonal distance distributions. For the data set with 0% substrate contribution, nine orthogonal outliers were removed from the IL class in the first step, and one orthogonal outlier was removed from the ALI group in the second step.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The data set developed with 20% substrate contribution had 86 orthogonal outliers removed from the IL class in the first step, six orthogonal outliers removed from the ALI group in step two, and eight orthogonal outliers removed from the ARG group in step two. One additional IL outlier (with 0% substrate contribution), was removed from the 20% substrate contribution data set but not removed from the 0% substrate contribution data set. The remaining 76 outliers removed from the 20% substrate contribution data set were electronically generated from those IL records that were also removed.

For each step of the multi-step classification scheme, the value of gamma was optimized. The optimized values, shown in columns one and two of Table 7 for the 0% and 20% substrate contribution models, were chosen based on the highest correct classification rates for the cross-validation test samples. For both data sets, the steps which classify a sample into a single ASTM class (AR, GAS, ISO, NA, PD, and NP) required a gamma value of 1 for optimal results. This indicates that the orthogonal distance is more important for correct classification in steps where the TIS for the classes are similar. The optimized γ is smaller for classification steps where the TIS for the classes are less similar (i.e., IL versus SUB and ALI versus ARG), reflecting an increased importance of the score distance. The optimization of gamma for the ISO/NA versus PD/NP classification step was insensitive to the magnitude of γ (range 98% - 99% correct). Given the similarity of the TIS for the combined ISO/NA and PD/NP groups, a γ of 0.8 was selected for this classification step.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 7: Values of γ used in the SIMCA method. The rows are labeled with the specific step of the classification scheme, while the middle and right columns show the γ values used.

	0% Substrate Contribution		20% Substrate Contribution
Step	γ		γ
IL, SUB	0.4		0.4
ALI, ARG	0.2		0.2
AR, GAS	1		1
ISO/NA, PD/NP	0.8		0.8
ISO, NA	1		1
PD, NP	1		1

As discussed above, in the first step of the classification scheme, samples were determined to be "positive" if IL was present and "negative" if IL was not present. This allows the designation of true positive, false positive, true negative and false negative rates for the first classification step. Table 8 shows the true positive rate (TPR), false positive rate (FPR), precision and accuracy values for the first step of the multi-step scheme, which involves classification into the IL and SUB classes. Table 8 entries represent results for the model developed with 0% substrate contribution and the numbers given in parentheses represent the values determined from the model developed with 20% substrate contribution. SIMCA based on the 20% substrate contribution data set gave very good results, with a 94.2% TPR and 5.1% FPR with greater than 94% precision and accuracy. The level of substrate contribution and the amount of remaining ignitable liquid are uncontrolled in fire debris samples containing an ignitable liquid. In order to account for these uncontrolled factors, the "Fire Debris" performance metrics are given for fire debris class designations based on the sample's proximity to the pour and the analyst's examination of the data, as discussed in the preceding sections. SIMCA detection of the presence of IL trace was in good agreement with the analyst, showing greater than 95% TPR; however, the FPR also increased to approximately 15%. This result

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

means that the analyst is slightly more conservative than SIMCA in determining a sample to be positive for IL residue. Nonetheless, given the extremely complicated nature of the problem, SIMCA performs well on the important step of determining if a sample is positive for IL residue.

Table 8: Performance metrics, shown as percentages, for the cross-validation test set and fire debris samples using SIMCA. Values for the 0% substrate contribution data set are shown and the 20% substrate contribution data set values are given in parentheses.

	Cross-Validation	Fire Debris - Pour	Fire Debris - Analyst
TPR*	94.0 (94.2)	81.0 (79.1)	96.7 (95.1)
FPR†	17.3 (5.1)	8.9 (8.9)	16.3 (15.0)
Precision	96.4 (95.4)	97.0 (96.9)	90.2 (90.7)
Accuracy	92.1 (94.5)	83.3 (81.8)	91.6 (91.1)

* True Positive Rate (TPR)
† False Positive Rate (FPR)

0% Substrate Contribution

Figure 21 shows confusion matrices of the classification results for the model developed with 0% substrate contribution. Results are shown for the TIS of cross-validation test samples and fire debris with class designations based on proximity to the pour. Each row is labeled with the designated class and the column headings represent the class assigned by the model. The entries in the table give the number of TIS assigned to each class. For example, out of 9,000 IL TIS, 8,454 were correctly assigned to IL, 543 were incorrectly assigned to SUB, and 3 were not assigned by the model. This gives a correct classification rate of 94.0% (corresponding to the true positive rate of 94.0% in Table 8). This same process is used to interpret the values in each confusion matrix in Figure 22.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Not Assigned	Total	% Correct
IL	8,454 (128)	543 (30)	3 (0)	9,000(158)	94.0 (81.0)
SUB	312 (4)	1,488 (41)	0 (0)	1,800 (45)	82.7 (91.1)
Total	8,766 (132)	2,031 (71)	3 (0)	10,800 (203)	92.1 (83.3)

	ALI	ARG	Not Assigned	Total	% Correct
ALI	4,676 (39)	24 (27)	0 (0)	4,700(66)	99.5 (59.1)
ARG	18 (0)	1,182 (82)	0 (0)	1,200 (82)	98.5 (100.0)
Total	4,686 (39)	1,214 (109)	0 (0)	5,900 (148)	99.3 (81.8)

	AR	GAS	Not Assigned	Total	% Correct
AR	486 (9)	0 (0)	14 (0)	500 (9)	100.0 (100.0)
GAS	43 (30)	752 (20)	5 (23)	800 (73)	94.6 (40.0)
Total	529 (39)	752 (20)	19 (23)	1,300 (82)	96.6 (49.2)

	ISO/NA	PD/NP	Not Assigned	Total	% Correct
ISO/NA	930 (10)	65 (3)	5 (6)	1,000 (19)	93.5 (76.9)
PD/NP	16 (0)	3,676 (27)	8 (20)	3,700 (47)	99.6 (100.0)
Total	946 (10)	3,741 (30)	13 (26)	4,700 (66)	98.3 (92.5)

	ISO	NA	Not Assigned	Total	% Correct
ISO	644 (7)	0 (0)	56 (5)	700 (12)	100.0 (100.0)
NA	0 (1)	300 (0)	0 (6)	300 (7)	100.0 (0.0)
Total	644 (8)	300 (0)	56 (11)	1,000 (19)	100.0 (87.5)

	PD	NP	Not Assigned	Total	% Correct
PD	3,256 (18)	20 (0)	24 (18)	3,300 (36)	99.4 (100.0)
NP	141 (7)	159 (0)	0 (4)	300 (11)	53.0 (0.0)
Total	3,397 (25)	179 (0)	24 (22)	3,600 (47)	95.5 (72.0)

Figure 21: Multi-step classification scheme with SIMCA cumulative results for cross-validation using 0% substrate contribution. The results for the fire debris samples, based on proximity to the pour, are shown in parentheses.

	IL	SUB	Not Assigned	Total	% Correct
IL	119	2	0	123	96.7
SUB	13	67	0	80	83.8
Total	132	71	0	203	91.6

	ALI	ARG	Not Assigned	Total	% Correct
ALI	39	11	0	50	78.0
ARG	0	66	0	66	100.0
Total	39	77	0	116	90.5

	AR	GAS	Not Assigned	Total	% Correct
AR	9	0	0	9	100.0
GAS	30	18	9	57	37.5
Total	39	18	9	66	47.4

	ISO/NA	PD/NP	Not Assigned	Total	% Correct
ISO/NA	10	3	5	18	76.9
PD/NP	0	27	5	32	100.0
Total	10	30	10	50	92.5

	ISO	NA	Not Assigned	Total	% Correct
ISO	7	0	4	11	100.0
NA	1	0	6	7	0.0
Total	8	0	10	18	87.5

	PD	NP	Not Assigned	Total	% Correct
PD	18	0	6	24	100.0
NP	7	0	1	8	0.0
Total	25	0	7	32	72.0

Figure 22: Confusion matrices for fire debris samples based on models with 0% substrate contribution. Designations for the fire debris samples are based on the analyst's review of the post-burn data.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

For the cross-validation test samples, a total correct classification rate of 92.1% was observed for a total of 10,800 IL/SUB TIS. When the model was tested on the TIS of fire debris samples, with class designations based on proximity to the pour, a correct classification rate of 83.3% was observed with the TIS of 30 IL-designated samples misclassifying as SUB. The TIS of four SUB-designated samples misclassified as IL and examination of the data did not reveal any clear reason for the error. When the fire debris designations were based on the analyst's examination of the data, a correct classification rate of 91.6% was observed. The increase in correct classification rate is attributed to the analyst's conservative approach, requiring a stronger IL residue to determine a sample positive for IL. The stronger IL contribution results in a higher percent correct classification rate using SIMCA.

The next step, classifying the TIS of an IL-designated sample as either ARG or ALI, gave a correct classification rate of 99.3% for the cross-validation test samples and 81.8% for fire debris with designations based on the proximity to the pour. A correct classification rate of 90.5% was observed when using the class designations based on the analyst's examination of the data.

Samples designated as ARG were further classified into their respective ASTM E1618-10 classes. The TIS for cross-validation test samples reflected a 96.6% correct classification rate, while the TIS for fire debris with designations based on proximity to the pour, performed much worse with an overall correct classification rate of 49.2%, which is somewhat misleading. All AR fire debris samples were correctly classified, while 60% of GAS fire debris samples assigned were misclassified as AR. This may be attributed to evaporation of the lighter, primarily aliphatic, gasoline components in the sample leaving the heavier, primarily aromatic, components. The post-burn TIS profile for gasoline samples changes significantly from the TIS

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

for fresh gasoline, leading to the misclassifications. Caution would be required when trying to classify a fire debris AR or GAS sample into the correct ASTM class using the SIMCA approach investigated here. When using designations for fire debris based on the analyst's examination of the data, a correct classification rate of 47.4% was observed. The AR samples were correctly assigned, while 63% of those GAS samples assigned were misclassified as AR. Notably, in a previous report, classification using QDA and LDA did not suffer from this difficulty (66).

The next step of the classification scheme assigns the TIS of samples designated as ALI to either the combined ISO/NA group or the combined PD/NP group. Cross-validation test samples for this step had high correct classification rates of 98.3% when predicting classifications of 4,700 TIS. Using class designations based on proximity to the pour for fire debris, a correct classification rate of 92.5% was observed. The correct classification rate was also 92.5% when class designations were based on the analyst's examination of the data.

The TIS for samples designated as members of the ISO/NA or PD/NP groups were separated into their respective ASTM classes. Classification of the TIS for cross-validation test samples into ISO or NA resulted in a correct classification rate of 100%, while fire debris, with designations based on proximity to the pour, was 87.5% correct. In the data set, seven of 12 ISO samples were assigned and all classified correctly, while one NA sample was incorrectly assigned, and the remaining six NA samples were unassigned. When using the designations based on the analyst's examination of the fire debris data, a correct classification rate of 87.5% was also observed. The pattern of correct classification was the same as described for class designation based on proximity to the pour. When assigning a sample to the PD or NP class, correct classification rates were 95.5% for the TIS of cross-validation test samples and 72.0% for fire debris, with designations based on proximity to the pour. In the data set, 18 of 36 PD

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

samples were assigned and all classified correctly, while seven NP samples were incorrectly assigned, and four were unassigned. A correct classification rate of 72.0% was also obtained when using the designations based on the analyst's examination of the data for fire debris. The pattern of correct classification was the same as described for class designation based on proximity to the pour.

20% Substrate Contribution

Figure 23 shows confusion matrices for classifications based on the model developed with up to 20% substrate contribution. Results are given for the TIS of cross-validation test samples and fire debris samples. For the first step of the classification scheme, a total correct classification rate of 94.5% was observed for the TIS of 189,300 cross-validation test samples, while the TIS of 203 fire debris samples, with designations based on the proximity to the pour, resulted in a correct classification rate of 81.8%. Classification of the fire debris samples resulted in the TIS of 33 IL-designated samples being misclassified as SUB. The TIS of four SUB-designated samples misclassified as IL; these four samples were also observed to misclassify when using the model developed with 0% substrate contribution. Using designations based on the analyst's examination of the fire debris data, a correct classification rate of 91.1% was observed. Including substrate contribution into the model data set did not significantly improve or decrease the rates of correctly classifying a sample as positive for IL residue.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

	IL	SUB	Not Assigned	Total	%Correct
IL	93,686 (125)	5,814 (33)	0 (0)	99,500 (158)	94.2 (79.1)
SUB	4,540 (4)	85,260 (41)	0 (0)	89,800 (45)	94.9 (91.1)
Total	98,226 (129)	91,074 (74)	0 (0)	189,300 (203)	94.5 (81.8)

	ALI	ARG	Not Assigned	Total	%Correct
ALI	51,380 (42)	220 (24)	0 (0)	51,600 (66)	99.6 (63.6)
ARG	171 (1)	13,329 (81)	0 (0)	13,500 (82)	98.7 (98.8)
Total	51,551 (43)	13,549 (105)	0 (0)	65,100 (148)	99.4 (83.1)

	AR	GAS	Not Assigned	Total	%Correct
AR	5,299 (9)	0 (0)	1 (0)	5,300 (9)	100.0 (100.0)
GAS	132 (37)	8,268 (14)	0 (22)	8,400 (73)	98.4 (27.5)
Total	5,431(46)	8,268 (14)	1 (22)	13,700 (82)	99.0 (38.3)

	ISO/NA	PD/NP	Not Assigned	Total	%Correct
ISO/NA	11,264 (13)	136 (3)	0 (3)	11,400 (19)	98.8 (81.3)
PD/NP	3 (0)	40,297 (35)	0 (12)	40,300 (47)	99.9 (100.0)
Total	11,267 (13)	40,433 (38)	0 (15)	51,700 (66)	99.7 (94.1)

	ISO	NA	Not Assigned	Total	%Correct
ISO	7,700 (9)	0 (0)	0 (3)	7,700 (12)	100.0 (100.0)
NA	0 (2)	3,700 (3)	0 (2)	3,700 (7)	100.0 (60.0)
Total	7,700 (11)	3,700 (3)	0 (5)	11,400 (19)	100.0 (85.7)

	PD	NP	Not Assigned	Total	%Correct
PD	36,166 (22)	534 (0)	0 (14)	36,700 (36)	98.5 (100.0)
NP	835 (10)	2,665 (1)	0 (0)	3,500 (11)	76.1 (9.1)
Total	37,001 (32)	3,199 (1)	0 (14)	40,200 (47)	96.6 (69.7)

Figure 23: Multi-step classification scheme with SIMCA cumulative results for cross-validation using 20% substrate contribution. The results for the fire debris samples, based on proximity to the pour, are shown in parentheses.

	IL	SUB	Not Assigned	Total	%Correct
IL	117	6	0	123	95.1
SUB	12	68	0	80	85.0
Total	129	74	0	203	91.1

	ALI	ARG	Not Assigned	Total	%Correct
ALI	41	9	0	50	82.0
ARG	0	66	0	66	100.0
Total	41	75	0	116	92.2

	AR	GAS	Not Assigned	Total	%Correct
AR	9	0	0	9	100.0
GAS	37	11	9	57	22.9
Total	46	11	9	66	35.1

	ISO/NA	PD/NP	Not Assigned	Total	%Correct
ISO/NA	12	3	3	18	80.0
PD/NP	0	28	4	32	100.0
Total	12	31	7	50	93.0

	ISO	NA	Not Assigned	Total	%Correct
ISO	8	0	3	11	100.0
NA	2	3	2	7	60.0
Total	10	3	5	18	84.6

	PD	NP	Not Assigned	Total	%Correct
PD	20	0	4	24	100.0
NP	8	0	0	8	0.0
Total	28	0	4	32	71.4

Figure 24: Confusion matrices for fire debris samples based on models with an upper limit of 20% substrate contribution. Designations for the fire debris samples are based on the analyst's review of the post-burn data.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

The next step classified IL-designated samples as belonging to the ALI or ARG group. For the TIS of cross-validation test samples, 99.4% classified correctly. The TIS of fire debris samples, with designations based on proximity to the pour, resulted in a correct classification rate of 83.1%. When designations based on the analyst's examination of the fire debris data were used, a correct classification rate of 92.2% was obtained. Including substrate contribution in the model data set did not significantly change the correct classification rates for this step in the classification scheme, and similar comments hold for the other steps in the classification, as discussed below.

When samples designated as ARG were assigned to their individual classes, a correct classification rate of 99.0% was observed for the TIS of the cross-validation test samples. The correct classification rate for the TIS of fire debris samples was 38.3%, with all AR samples classifying correctly and 73% of GAS samples that were assigned misclassified as AR. This can also be attributed to weathering as previously discussed for the data set with 0% substrate contribution. Using the designations based on the analyst's examination of the fire debris data, the correct classification rate was 35.1%.

Samples designated as ALI were separated into the combined ISO/NA group and the combined PD/NP group. When predicting the classification of the TIS of 51,700 cross-validation test samples, a correct classification rate of 99.7% was obtained. For the TIS of fire debris samples, with designations based on proximity to the pour, the correct classification rate was 94.1%. A correct classification rate of 93.0% was obtained when using the designations based on the analyst's examination of the fire debris data.

The classification of the TIS of cross-validation test samples into the ISO or NA classes resulted in a correct classification rate of 100.0%. The TIS of fire debris samples, with

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

designations based on proximity to the pour, had a correct classification rate of 85.7%. In the data set, nine of 12 ISO samples were assigned and all classified correctly, while three NA samples were correctly assigned, two were incorrectly assigned, and two were unassigned. A correct classification rate of 84.6% was obtained when the designations were based on the analyst's examination of the fire debris data. The pattern of correct classification was similar to that described for class designation based on proximity to the pour. When assigning the TIS of cross-validation test samples to either the PD or NP class, a correct classification rate of 96.6% was observed. The TIS of fire debris samples, with designations based on proximity to the pour, gave a correct classification rate of 69.7%. In the data set, 22 of 36 PD samples were assigned and all classified correctly, while 10 of the 11 NP samples were incorrectly assigned. A correct classification rate of 71.4% was obtained when using the designations based on the analyst. The pattern of correct classification was similar to that described for class designation based on proximity to the pour.

Results for Different Normalization Methods

Summed to One

Figure 25 gives the summarized results for LDA, QDA, and SIMCA using the "summed to one" normalization method. Each leg of the graph represents the percent correct classification for each step in the classification scheme. The cross-validation test set results are shown as blue, filled diamonds connected by a solid line and the fire debris results, with designations based on the proximity to the pour, are represented by green, filled squares and a broken line. The fire debris samples with designations based on the analyst's review of the post-burn data are shown

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

as red, filled triangles connected by a broken line. A perfect classification system would be reflected in a graph with both the cross-validation and fire debris points at 100% correct (i.e., all points on the outermost periphery of the graph). These same guidelines can be used to interpret the graphs in Figure 26 and Figure 27. In Figure 25, SIMCA is observed to perform better than LDA and QDA for cross validation, but not for fire debris data. The lowest correct classification rates for the LDA and QDA models occur when discriminating between the ISO and NA fire debris samples. For the SIMCA model, the lowest correct classification rates were obtained when discriminating between the AR and GAS fire debris samples. This normalization is a linear transformation of the intensities of each sample. In samples with fewer chromatographic peaks, noise and baseline contributions will have a greater influence on inter-TIS comparisons.

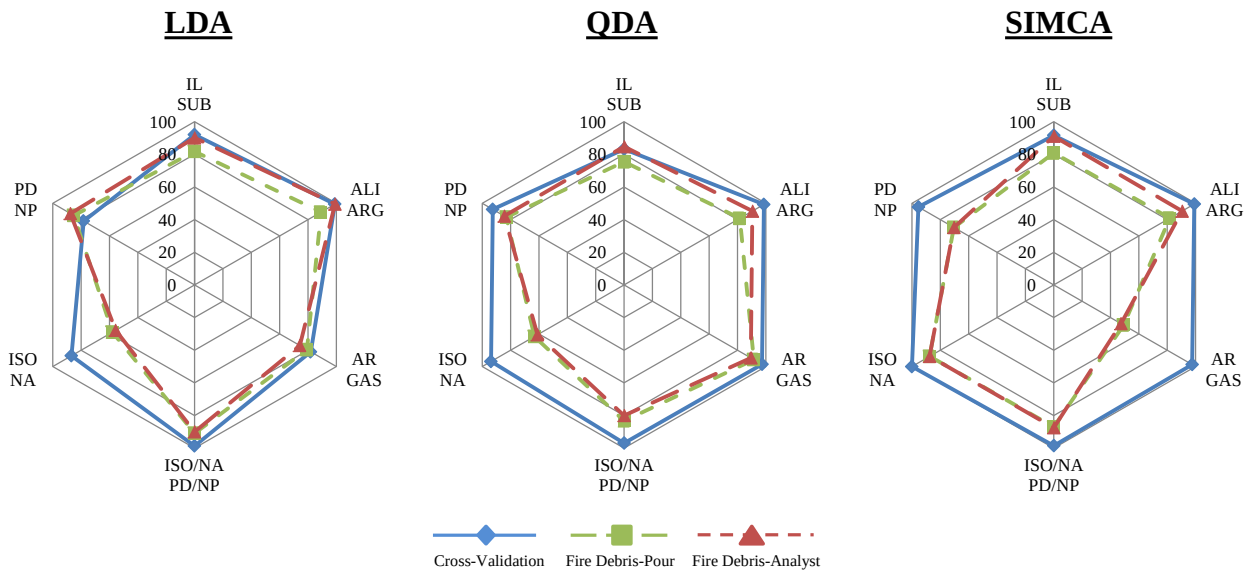


Figure 25: Correct classification rates - "summed to one" normalization method.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Base Peak

The summarized results using the "base peak" normalization method are given in Figure 26. The SIMCA method is again observed to perform the best overall. Compared to the previous normalization method, the fire debris samples showed a dramatic increase in correct classification rates. The results were nearly perfect for all steps except the steps which discriminate the fire debris samples in the IL and SUB classes and the PD and NP classes. Increased correct classification rates are observed because the fire debris samples that previously misclassified, primarily the GAS-designated samples, were not assigned using this normalization method and set of γ values. This data normalization method results in more conservative assignments by the SIMCA classification method. This normalization method is also a linear transformation, but inter-TIS comparisons for samples containing a small number of chromatographic peaks are less influenced by baseline and noise. This normalization is commonly used for mass spectral normalization.

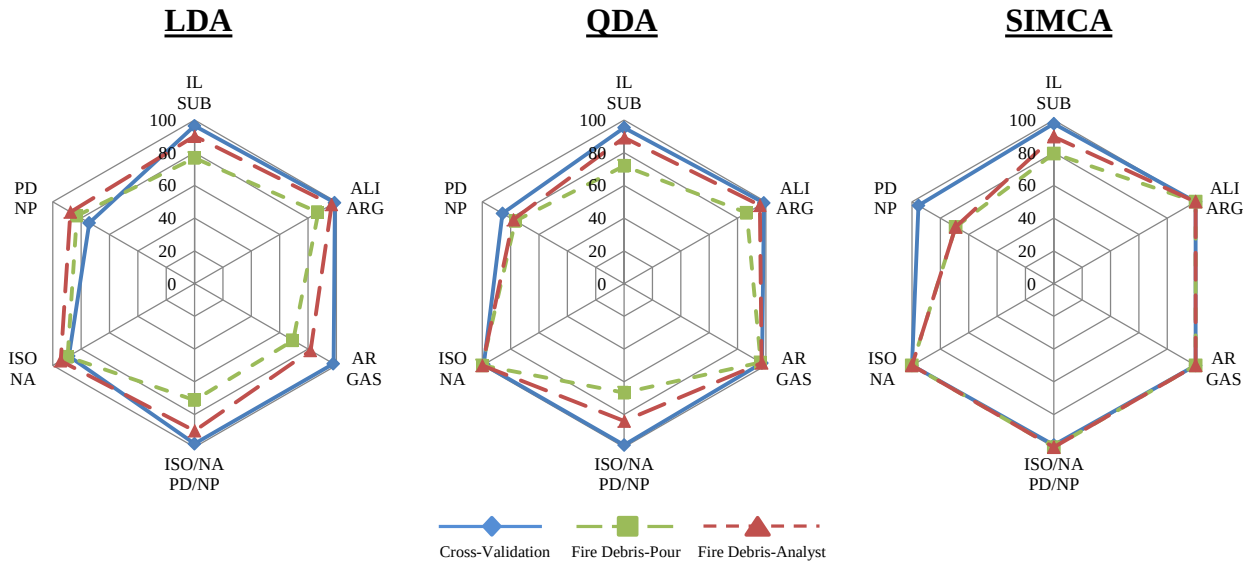


Figure 26: Correct classification rates - "base peak" normalization method.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Unit Vector

The results obtained for the "unit vector" normalization are shown in Figure 27. Again, SIMCA performed the best overall, and very high correct classification rates were obtained for the fire debris samples. As discussed for the results with the "base peak" normalization, the majority of the fire debris samples were not assigned; therefore, this data normalization method also resulted in more conservative assignments by the SIMCA method. Although class information is not provided for the samples that are unassigned, which could be seen as a potential downfall of the SIMCA method, a misclassification is also prevented. Similar results observed for the "base peak" and "unit vector" normalization methods indicate that the increase in correct classification rate is based on the set of γ values or normalization method. The same γ values were investigated for the "summed to one" normalization methods, but the correct classification rates did not improve. This normalization is a linear transformation that converts each TIS into a vector of unit length. In this normalization, the baseline and noise also influence the inter-TIS comparisons, but this influence is less than for the sum-to-one method and greater than the base-peak method.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

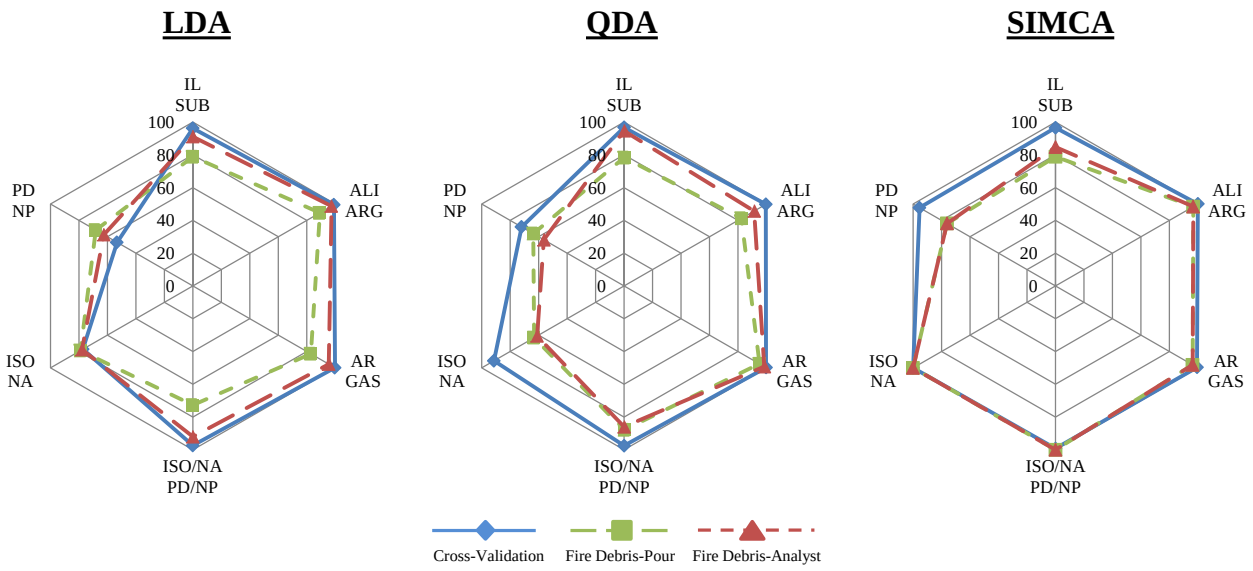


Figure 27: Correct classification rates - "unit vector" normalization method.

The results demonstrate that the correct classifications rates are influenced by the normalization methods used. For the fire debris samples, low correct classification rates were observed when assigning samples to the PD or NP classes using all three normalization methods. The low rates were due to the NP-designated samples misclassifying as PD. When the data were examined, it was determined that the TIS for samples in these classes are very similar. The TICs, however, are not as similar and could be used to determine the sample's class assignment.

The true positive and false positive rates are designated as TPR and FPR, respectively, in Table 9. The precision associated with each data set is given as well. These values are expressed as percentages, and the precision represents the number of IL-designated samples that was assigned to the IL class out of the total number of samples assigned to the IL class. These metrics were calculated for each normalization method using each classification method. For example, the "summed to one" normalization method and LDA classification method resulted in a 95.5% true positive rate for the cross-validation test set.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Table 9: Classifier Performance Metrics for Data Normalization Methods

Classification Method	Normalization Method	Cross-Validation			Fire Debris-Pour			Fire Debris-Analyst		
		TPR	FPR	Precision	TPR	FPR	Precision	TPR	FPR	Precision
LDA	Summed to One	95.5	25.4	95.0	81.0	15.6	94.8	96.8	20.0	88.1
	Base Peak	98.2	12.7	97.5	72.8	8.9	96.6	90.2	10.0	93.3
	Unit Vector	97.9	12.6	97.5	74.1	4.4	98.3	91.1	8.8	94.1
QDA	Summed to One	81.7	9.1	97.9	70.9	8.9	96.6	84.6	15.0	89.7
	Base Peak	96.2	9.3	98.1	64.6	2.2	99.0	82.9	1.3	99.0
	Unit Vector	98.3	11.2	97.8	72.8	2.2	99.1	92.7	2.5	98.3
SIMCA	Summed to One	93.0	15.3	96.8	77.2	6.7	97.6	93.5	12.5	92.0
	Base Peak	97.5	1.1	99.8	78.1	15.4	94.7	95.2	18.3	88.5
	Unit Vector	96.4	2.8	99.5	76.7	14.0	94.9	88.7	21.6	86.4

Classification by the SIMCA method and normalization by the "base peak" method resulted in the highest true positive rates and lowest false positive rates for the cross-validation test set. These methods did not perform as well for the fire debris samples and reflected a false positive rate that was nearly double the rate obtained using the "summed to one" method. When considering the QDA classification method, low false positive rates were obtained for the model and fire debris samples when using the "base peak" method. The true positive rate was also low using these methods on the fire debris samples with designations based on the proximity to the pour; however, the rates were improved when using the designations based on the analyst. High precision values for both fire debris data sets indicate that the misclassifications resulted from IL-designated samples assigning to the SUB class. When considering the fire debris data set with designations based on the analyst, the true positive rates and overall correct classification rate improved. This results due to the more-conservative designations made by the analyst in

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

determining if an IL was present in the post-burn sample. When LDA was used for classification, the lowest false positive rates were obtained with "unit vector" normalization of the data. The false positive rates were higher for the cross-validation test set than for the fire debris samples.

Training Results

An international group of 35 fire debris analysts volunteered for the course and only 25 (71%) participants (9 U.S. and 16 European) finished the course online. A second group of 15 U.S. fire debris analysts enrolled to participate in the face-to-face training. A total of 14 (93%) of the 15 participants were actually trained face-to-face. This difference in completion rates (71% versus 93%, an effect size of 0.61) is not statistically significant (α=0.13) at a power of 0.8, which is often considered satisfactory for a small sample size. The average score on the final examination in the course was 92% for the online participants and 90% for the face-to-face participants. This difference is not statistically significant and an indication that the methods can be trained either online or face-to-face. The demographic data for both sets of participants (online and face-to-face were fairly comparable (see Appendix II at the end of this report), with the most significant differences being that the face-to-face participants were older, self-reported to be more proficient in English and had more years of experience in fire debris analysis.

IV. Conclusions

Discussion of Findings

The findings from this research demonstrate that it is possible to develop chemometric methods that provide reliable error rates for fire debris analysis and that it is possible to transition these

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

methods to operational forensic laboratories by online training. The challenge of inter-laboratory variability of chromatographic retention time was overcome by the use of the total ion spectrum, or TIS. The TIS was further investigated and shown to provide a high degree of discrimination between individual ignitable liquids and classes of ignitable liquids as defined under ASTM E1618 for well-characterized classes. In some cases, the mass spectral data alone cannot provide a high degree of discrimination between ASTM classes (i.e., PD vs. NP and NA vs. ISO). Furthermore, the miscellaneous category, which is devoid of any true class characteristic is problematic.

The methods of LDA, QDA and SIMCA were shown to provide relatively good discrimination of ignitable liquid and pyrolysis samples in cross validation and when applied to the analysis of fire debris samples. The best overall performance was observed for QDA classification using the "base peak" method. In subsequent steps of the classification into ASTM classes for those samples that were classified as positive for ILR, the LDA and QDA performance was varied depending on the data normalization method applied prior to model development. Nonetheless the results for most classifications were greater than 80%, with some problems observed for discrimination between highly similar ASTM classes (i.e., PD vs. NP and NA vs. ISO). The results of these classifications are considered less important forensically than the classification of a sample as positive or negative for ILR. The correct classification rates for samples that were positive or negative for ILR are encouraging and reflect the possibilities for introducing statistical methods into the analysis of fire debris in order to bring this area of forensic analysis into closer alignment with the Daubert standard.

Finally, this research demonstrates that it is possible to transfer basic research methods to operational laboratories through training. One difficulty is in knowing when to transfer the

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

methods. Certainly the methods should not be transferred or trained until they have been published under peer-review. Once published, the methods should be trained to operational laboratory personnel; otherwise, the methods will remain in the literature and unused, never achieving general acceptance.

Implications for Policy and Practice

The results of this research provide the first large-scale demonstration of statistically reliable classification rates for fire debris as positive or negative for ignitable liquid residue. Fire debris analysis methods with known error rates meet the Daubert requirements and may someday be required under existing rules of evidence. We believe that future research will result in methods that provide better correct classification rates, nonetheless, the current study provides a basis for future work and demonstrates that statistical methods may be applied to system as complex as fire debris analysis, resulting in known and defensible error rates. The research will help to drive policy change to improve forensic fire debris analysis. The methods have been transferred to operational laboratories by both online and face-to-face training. Successfully training these methods online sets precedence for future research and demonstrates one route to influencing laboratory practice.

Implications for Further Research

Although the results of this research are viewed to make a positive step forward in the search for reliable methods for assessing fire debris as positive or negative for the presence of ignitable liquid residue, additional research is needed. Of particular importance are (1) the need for a greater sampling of substrate pyrolysis samples, (2) the need to examine statistical methods that are less dependent on normal distributions of the data sets, (3) methods for incorporating both the chromatographic and mass spectral data in the statistical assessment, and (4) new Bayesian

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

statistical approaches that calculate likelihood ratios (i.e., the probability that a sample contains ignitable liquid residue relative to the probability that it does not).

V. References

1. ASTM. E1386 -10 Standard Practice for Separation of Ignitable Liquid Residues from Fire Debris Samples by Solvent Extraction. West Conshohocken, PA: ASTM International; 2010.
2. ASTM. E1388-12 Standard Practice for Sampling of Headspace Vapors from Fire Debris Samples. West Conshohocken, PA: ASTM International; 2012.
3. ASTM. E1412-12 Standard Practice for Separation of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Activated Charcoal. West Conshohocken, PA: ASTM International; 2012.
4. ASTM. E1413-07 Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Dynamic Headspace Concentration. West Conshohocken, PA: ASTM International; 2007.
5. ASTM. E2154-01(2008) Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Solid Phase Microextraction (SPME). West Conshohocken, PA: ASTM International; 2008.
6. ASTM. E1618-11 Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. West Conshohocken, PA: ASTM International; 2011.
7. Forensic Science Handbook. 2nd ed. Upper Saddle River, New Jersey: Prantice Hall, 2002.
8. Sigman ME, Williams MR, Castelbuono JA, Colca JG, Clark CD. Ignitable Liquid Classification and Identification Using the Summed-Ion Mass Spectrum. Instrum Sci Technol. 2008;36(4):375-93.
9. Grotch SL. Matching of mass spectra when peak height is encoded to one bit. Anal Chem. 1970;42(11):1214-22.
10. Doble P, Sandercock M, Du Pasquier E, Petocz P, Roux C, Dawson M. Classification of premium and regular gasoline by gas chromatography/mass spectrometry, principal component analysis and artificial neural networks. Forensic Sci Int. 2003 Mar 12;132(1):26-39.
11. Bertsch W, Sellers CS, Babin K, Holzer G. Automation in the chemical analysis of suspect arson samples by GC/MS. A systematic approach. Journal of High Resolution Chromatography. 1988;11(11):815-9.
12. Keto RO, Wineman PL. Detection of petroleum-based accelerants in fire debris by target compound gas chromatography/mass spectrometry. Analytical Chemistry. 1991 1991/09/01;63(18):1964-71.
13. Lennard CJ, Tristan Rochaix V, Margot P, Huber K. A GC–MS database of target compound chromatograms for the identification of arson accelerants. Science & Justice. 1995 1//;35(1):19-30.
14. The Ignitable Liquids Reference Database can be found at the universal resource locator; http://ilrc.ucf.edu. [cited April 26, 2013]; Available from.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

15. Murphy DM, Middlebrook AM, Warshawsky M. Cluster Analysis of Data from the Particle Analysis by Laser Mass Spectrometry (PALMS) Instrument. Aerosol Sci and Technol. 2003;37(4):382-91.
16. Lien-Chin C, Yu-Cheng L, Chi-Wei L, Tseng VS, editors. A two-phase clustering approach for peak alignment in mining mass spectrometry data. Bioinformatics and Biomedicine Workshop, 2009 BIBMW 2009 IEEE International Conference on; 2009 1-4 Nov. 2009.
17. Heller SR, Chang CL, Chu KC. Interpretation of mass spectrometry data using cluster analysis. Alkyl thiolesters. Anal Chem. 1974;46(7):951-2.
18. Romesburg HC. Cluster analysis for researchers / H. Charles Romesburg. Malabar, Fla: Robert E. Krieger Pub. Co, 1990.
19. Sharaf MA, Illman DL, Kowalski BR. Chemometrics. New York: John Wiley & Sons, Inc., 1986.
20. Tan PN, Kumar V, Steinbach M. Introduction to data mining. 1st ed. ed. Boston: Pearson Addison Wesley, 2005.
21. Mat Desa WNS, Nic Daéid N, Ismail D, Savage K. Application of unsupervised chemometric analysis and self-organizing feature map (SOFM) for the classification of lighter fuels. Anal Chem. 2010;82(15):6395-400.
22. Wold S. Chemometrics: What do we mean with it, and what do we want from it? Chemometrics and Intelligent Laboratory Systems. 1995;30:109-15.
23. Waddell EE, Williams MR, Sigman ME. Progress Toward the Determination of Correct Classification Rates in Fire Debris Analysis II: Utilizing Soft Independent Modeling of Class Analogy (SIMCA). Journal of Forensic Sciences. 2014:n/a-n/a.
24. Fei BKL, Eloff JHP, Olivier MS, Venter HS. The use of self-organising maps for anomalous behavior detection in a digital investigation. Forensic Science International. 2006;162:33-7.
25. Kohonen T. Self-Organizing Maps. 3rd ed. New York: Sprinter, 2001.
26. Kohonen T. The Self-Organizing Map. Proceedings of the IEEE. 1990;79(9):1464-80.
27. Koua EL, editor. Using self-organizing maps for information visualization and knowledge discovery in complex geospatial datasets. Proceedings of the 21st International Cartographic Conference: 'Cartographic Renaissance'; 2003 10-16 August 2003; Durban, Sourth Africa.
28. Klobucar D, Subasic M. Using self-organizing maps in the visualization and analysis of forest inventory. iForest. 2012;5:216-23.
29. Borges C, Gómez-Carracedo MP, Andrade JM, Duarte MF, Biscaya JL, Aires-de-Sousa J. Geographical classification of weathered crude oil samples with unsupervised self-organizing maps and a consensus criterion. Chemometrics and Intelligent Laboratory Systems. 2010;101(1):43-55.
30. Beale MH, Hagan MT, Demuth HB. Neural Network Toolbox(TM) 7 User's Guide. Journal [serial on the Internet]. 2010 Date.
31. Goodacre R, Pygall J, Kell DB. Plant seed classification using pyrolysis mass spectrometry with unsupervised learning: The application of auto-associative and Kohonen artificial neural networks. Chemometrics and Intelligent Laboratory Systems. 1996;34:69-83.
32. Daszykowski M, Walczak B, Massart DL. Projection methods in chemistry. Chemometrics and Intelligent Laboratory Systems. 2003;65:97-112.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

33. Boishebert Vd, Giraudel J-L, Montury M. Characterization of strawberry varieties by SPME-GC-MS and Kohonen self-organizing map. Chemometrics and Intelligent Laboratory Systems. 2006;80:13-23.
34. Fonseca AM, Biscaya JL, Aires-de-Sousa J, Lobo AM. Geographical classification of crude oils by Kohonen self-organizing maps. Analytica Chimica Acta. 2006;556(2):374-82.
35. Mat Desa WNS, Ismail D, Nic Daéid N. Classification and source determination of medium petroleum distillates by a chemometric and artificial neural networks: a self organizing feature approach. Anal Chem. 2011;83(7745-7754).
36. Tan B, Hardy JK, Snavely RE. Accelerant classification by gas chromatography/mass spectrometry and multivariate pattern recognition. Anal Chim Acta. 2000;422:37-46.
37. Hupp AM, Marshall LJ, Campbell DI, Smith RW, McGuffin VL. Chemometric analysis of diesel fuel for forensic and environmental applications. Analytica Chimica Acta. 2008;606:159-71.
38. Borusiewicz R, Zadora G, Zieba-Palus J. Application of head-space analysis with passive adsorption for forensic purposes in the automated thermal desorption-gas chromatography–mass spectrometry system. Chromatographia. 2004;60:S133-42.
39. DeHaan JD. Kirk's Fire Investigation. 4th ed. Upper Saddle River, New Jersey: Prentice-Hall, 1997.
40. Kelly RL, Martz RM. Accelerant Identification in Fire Debris by Gas Chromatography/Mass Spectrometry Techniques. Journal of Forensic Sciences. 1984;29(3):714-22.
41. Neural network PC tools : a practical guide / edited by Russell C. Eberhart and Roy W. Dobbins ; with a foreword by Bernard Widrow. San Diego: Academic Press, 1990.
42. Sandercock PML, Du Pasquier E. Chemical fingerprinting of unevaporated automotive gasoline samples. Forensic Sci Int. 2003;134:1-10.
43. Sandercock PML, Du Pasquier E. Chemical fingerprinting of gasoline: 2. Comparison of unevaporated and evaporated automotive gasoline samples. Forensic Science International. 2004;140(1):43-59.
44. Petraco NDK, Gil M, Pizzola PA, Kubic TA. Statistical Discrimination of Liquid Gasoline Samples from Casework. Journal of Forensic Sciences. 2008;53(5):1092-101.
45. Johnson KJ, Synovec RE. Pattern recognition of jet fuels: comprehensive GC×GC with ANOVA-based feature selection and principal component analysis. Chemometrics and Intelligent Laboratory Systems. 2002 1/28/;60(1–2):225-37.
46. Aitken CGG, Lucy D. Evaluation of trace evidence in the form of multivariate data. Journal of the Royal Statistical Society: Series C (Applied Statistics). 2004;53(1):109-22.
47. Sigman ME, Williams MR. Covariance mapping in the analysis of ignitable liquids by gas chromatography/mass spectrometry. Anal Chem. 2006 Mar 1;78(5):1713-8.
48. Balabin RM, Safieva RZ, Lomakina EI. Wavelet neural network (WNN) approach for calibration model building based on gasoline near infrared (NIR) spectra. Chemometrics and Intelligent Laboratory Systems. 2008;93:58-62.
49. Balabin RM, Safieva RZ, Lomakina EI. Gasoline classification using near infrared (NIR) spectroscopy data: Comparison of multivariate techniques. Analytica Chimica Acta. 2010;671:27-35.
50. Barshick S-A. Analysis of Accelerants and Fire Debris using Aroma Detection Technology. Journal of Forensic Sciences. 1998;43(2):284-93.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

51. González-Rodríguez J, Sissons N, Robinson S. Fire debris analysis by Raman spectroscopy and chemometrics. Journal of Analytical and Applied Pyrolysis. 2011;91(1):210-8.
52. Lu Y, Harrington PB. Forensic Application of Gas Chromatography–Differential Mobility Spectrometry with Two-Way Classification of Ignitable Liquids from Fire Debris. Analytical Chemistry. 2007 2007/09/01;79(17):6752-9.
53. Lu Y, Chen P, Harrington P. Comparison of differential mobility spectrometry and mass spectrometry for gas chromatographic detection of ignitable liquids from fire debris using projected difference resolution. Anal Bioanal Chem. 2009 2009/08/01;394(8):2061-7.
54. Duda RO, Hart PE, Stork DG. Pattern Classification. 2 ed: Wiley, 2001.
55. Varmuza K, Filzmoser P. Introduction to Multivariate Statistical Analysis in Chemometrics. Boca Raton, FL: Taylor & Francis Group, 2009.
56. Tominaga Y. Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN. Chemometrics and Intelligent Laboratory Systems. 1999;49(1):105-15.
57. Svensson O, Josefson M, Langkilde FW. Classification of Chemically Modified Celluloses Using a Near-Infrared Spectrometer and Soft Independent Modeling of Class Analogies. Appl Spectrosc. 1997;51(12):1826-35.
58. Vanden Branden K, Hubert M. Robust classification in high dimensions based on the SIMCA Method. Chemometrics and Intelligent Laboratory Systems. 2005;79:10-21.
59. Meza-Márquez OG, Gallardo-Velázquez T, Osorio-Revilla G. Application of mid-infrared spectroscopy with multivariate analysis and soft independent modeling of class analogies (SIMCA) for the detection of adulterants in minced beef. Meat Science. 2010;86(2):511-9.
60. Odunsi K, Wollman RM, Ambrosone CB, Hutson A, McCann SE, Tammela J, et al. Detection of epithelial ovarian cancer using 1H-NMR-based metabonomics. International Journal of Cancer. 2005;113(5):782-8.
61. Khanmohammadi M, Garmarudi A, Ghasemi K, Jaliseh H, Kaviani A. Diagnosis of colon cancer by attenuated total reflectance-fourier transform infrared microspectroscopy and soft independent modeling of class analogy. Medical Oncology. 2009;26(3):292-7.
62. De Maesschalck R, Candolfi A, Massart DL, Heuerding S. Decision criteria for soft independent modelling of class analogy applied to near infrared data. Chemometrics and Intelligent Laboratory Systems. 1999;47:65-77.
63. Gemperline PJ, Webber LD, Cox FO. Raw materials testing using soft independent modeling of class analogy analysis of near-infrared reflectance spectra. Analytical Chemistry. 1989 1989/01/01;61(2):138-44.
64. Smidt E, Meissl K, Schwanninger M, Lechner P. Classification of waste materials using Fourier transform infrared spectroscopy and soft independent modeling of class analogy. Waste Management. 2008;28(10):1699-710.
65. Wold S. Pattern Recognition by Means of Disjoint Principal Components Models. Pattern Recognition. 1975;8:127-39.
66. Waddell EE, Song ET, Rinke CN, Williams MR, Sigman ME. Progress toward the determination of correct classification rates in fire debris analysis. J Forensic Sci. 2013 58(4):887-96.
67. The Substrate Reference Database can be found at the universal resource locator; http://ilrc.ucf.edu/substrate/. [cited April 26, 2013]; Available from.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

68. ASTM. E1412-07 Standard practice for separation of ignitable liquid residues from fire debris samples by passive ehadspace concentration with activated charcoal. West Conshohocken, PA ASTM International; 2007.
69. Williams MR, Sigman ME, Lewis J, Pitan KM. Combined target factor analysis and Bayesian soft-classification of interference-contaminated samples: Forensic Fire Debris Analysis. Forensic Sci Int. 2012;222:373-86.
70. McHugh KM. Determining the presence of an ignitable liquid residue in fire debris samples utilizing target factor analysis [thesis]. Orlando (FL): University of Central Florida, 2010.
71. Lewis JN. The application of chemometrics to the detection and classification of ignitable liquids in fire debris using the total ion spectrum [thesis]. Orlando (FL): University of Central Florida, 2011.
72. ASTM. E1618-10 Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. West Conshohocken, PA: ASTM International; 2010.
73. Bar-Joseph Z, Gifford DK, Jaakkola TS. Fast Optimal Leaf Ordering for Hierarchical Clustering. Bioinformatics. 2001 June 1, 2001;17(suppl 1):S22-S9.
74. Team RDC. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria; 2011.
75. Lucas A. amap: Another Multidimensional Analysis Package. R package version 0.8-7 ed. Vienna, Austria; 2011.
76. Wang J, Plataniotis KN, Lu J, Venetsanopoulos AN. Kernel quadratic discriminant analysis for small sample size problem. Pattern Recognition. 2008;41(5):1528-38.
77. Bensmail H, Celeux G. Regularized Gaussian Discriminant Analysis Through Eigenvalue Decomposition. Journal of the American Statistical Association. 1996;91(436):1743-8.
78. Krzanowski WJ. Principles of Multivariate Analysis: A User's Perspective. New York: Oxford University Press, Inc., 1988.
79. Fawcett T. An introduction to ROC analysis. Pattern Recognition Letters. 2006;27:861-74.
80. Naenna T, Embrechts MJ, Szymanski B, Sternickel K, editors. An expert system for cardiology diagnosis in health care systems. Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference; 2004 December 12-15; Gold Coast, Australia
81. Simula O, Ahola J, Alhoniemi E, Himberg J, Vesanto J. Self-organizing map in analysis of large-scale industrial systems. In: Kaski S, Oja E, editors. Kohonen maps. 1st ed. ed. Amsterdam ; New York: Elsevier; 1999;375-87.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

VI. Dissemination of Research Findings

Publications and Thesis

Waddell, Erin E.; Song, Emma T.; Rinke, Caitlin N.; Williams, Mary R.; Sigman, Michael E. "Progress Toward the Determination of Correct Classification Rates in Fire Debris Analysis." *J Forensic Sci*. 2013;58:887-96.

Waddell, Erin E.; Williams, Mary R.; Sigman, Michael E. "Progress Toward the Determination of Correct Classification Rates in Fire Debris Analysis II: Utilizing Soft Independent Modeling of Class Analogy (SIMCA)." *J Forensic Sci*. 2014. doi: 10.1111/1556-4029.12417.

Waddell, Erin E.; Frisch-Daiello, Jessica L.; Williams, Mary R.; Sigman, Michael E. "Hierarchical Cluster Analysis of Ignitable Liquids Based on the Total Ion Spectrum." *J Forensic Sci*. (In-press).

Frisch-Daiello, Jessica L.; Williams, Mary R.; Waddell, Erin E.; Sigman, Michael E. "Application of Self-Organizing Feature Maps to Analyze the Relationships between Ignitable Liquids and Selected Mass Spectral Ions." *Forensic Sci Int*. 2014:236:84-9.

CHEMOMETRIC APPLICATIONS TO A COMPLEX CLASSIFICATION PROBLEM: FORENSIC FIRE DEBRIS ANALYSIS, ERIN ELIZABETH WADDELL, PhD DISSERTATION, UNIVERSITY OF CENTRAL FLORIDA, 2013

Patents

None

Presentations

Waddell, Erin E.; Williams, Mary R.; Sigman, Michael E. "Comparison of SIMCA with QDA and LDA for the Identification and Classification of Ignitable Liquid Residues." Florida Annual Meeting and Exposition. Palm Harbor, FL. May 10, 2013. Oral Presentation.

Waddell, Erin E.; Williams, Mary R.; Sigman, Michael E. "Comparison of SIMCA with QDA and LDA for the Identification and Classification of Ignitable Liquid Residues." American Academy of Forensic Sciences Conference. Washington, DC. February 19-22, 2013. Oral Presentation.

Waddell, Erin E.; Song, Emma T.; Rinke, Caitlin N.; Williams, Mary R.; Sigman, Michael E. "ASTM Classification of Ignitable Liquids and Residues by Chemometric Techniques."

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

European Academy of Forensic Science Conference. The Hague, Netherlands. August 20-24, 2012. Poster Presentation.

Waddell, Erin E.; Song, Emma T.; Rinke, Caitlin N.; Williams, Mary R.; Sigman, Michael E. "Chemometric-Assisted Detection and Classification of Ignitable Liquids in Fire Debris." Florida Annual Meeting and Exposition. Palm Harbor, FL. May 18-19, 2012. Oral Presentation.

Waddell, Erin E.; Song, Emma T.; Rinke, Caitlin N.; Williams, Mary R.; Sigman, Michael E. "Chemometric-Assisted Detection and Classification of Ignitable Liquids in Fire Debris." American Academy of Forensic Sciences. Atlanta, GA. February 22-25, 2012. Poster Presentation.

Sigman, Michael E.; Williams, Mary R.; Rinke, Caitlin N.; Waddell, Erin E.; Song, Emma T.; Lewis, Jennifer. "Error Rates for Assignment of Ignitable Liquid ASTM Classification by Automated Methods." International Conference on Forensic Inference and Statistics. Seattle, WA. July 19-21, 2011. Poster Presentation.

Sigman, Michael E.; Williams, Mary R.; Lewis, Jennifer; Waddell, Erin E; Rinke, Caitlin N. "Advancing Fire Debris Analysis Through Chemometrics." Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy. Atlanta, GA. March 13-18, 2011. Oral Presentation.

Waddell, Erin E.; Rinke, Caitlin N.; Williams, Mary R.; Sigman, Michael E. "Chemometric Assisted Detection and Classification of Ignitable Liquids in Fire Debris." American Academy of Forensic Sciences. Chicago, IL. February 23-26, 2011. Poster Presentation.

Appendices

Appendix I: Training Outline

Section 1: AMDIS Students are guided through the process of downloading and installing AMDIS on their computer. AMDIS is gas chromatography-mass spectrometry analysis software provided free of charge. The software allows students to read data from most commercial instruments. Students in this course were taught to use AMDIS to load example data files and extract the TIS for further analysis.

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Section 2: R The statistical methods developed under this research were produced using statistical software from the R Project. This is open-source software that can be downloaded free of charge from the internet. The use of R allows students in operational laboratories to implement the methods developed without the added expense of many commercial software packages. Students were taught to download and install R on their computers. They were also taught to install computational packages need to complete the training.

Section 3: TIS Students were given a refresher or introduction to the total ion spectrum (TIS) and the utility of the TIS for laboratory-independent representation of GC-MS data from fire debris, ignitable liquids and other complex systems. Students were taught to calculate and save the TIS from example data sets.

Section 4: Discriminant Analysis Students were taught the basic background information on data dimension reduction by principal components analysis (PCA) and the use of linear and quadratic discriminant analysis (LDA and QDA) to discriminate between classes using the PCA scores.

Section 5: Putting it All Together Students were taught to download and install an R package to calculate a QDA model for IL and SUB TIS data incorporated in the package. They were then taught to open GC-MS data sets with AMDIS, extract the TIS and save it to their local computer. The saved TIS file was then projected into the PCA space and the scores projected into the QDA model to provide a classification of the sample as positive or negative for ignitable liquid residue. The cross-validation classification errors were provided to the student in a confusion matrix and the student was taught to interpret the results.

Final Exam Students were given a final exam to assess their comprehension of the materials presented in the course. The exam included both theoretical and practical problems.

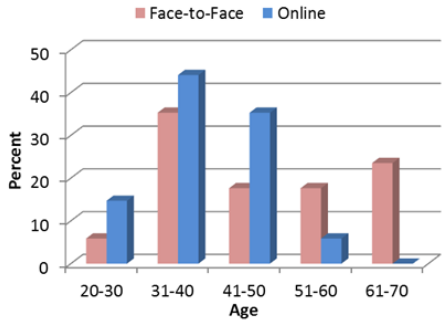
This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

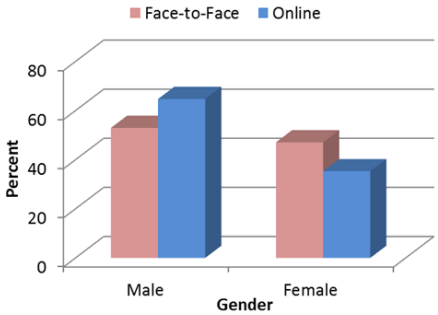
End of Course Survey Students were given an end-of-course survey to collect their feedback on the course.

Appendix II: Demographic data for online and face-to-face applicants for the training

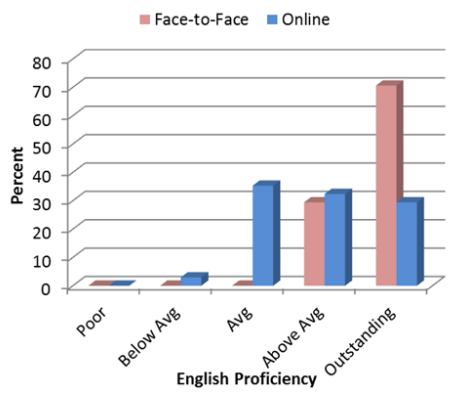
Thirty-five applicants for the online training and 17 requested face-to-face training. All participants were required to return a questionnaire containing 10 questions. The following graphs compare the online and face-to-face participants as self-reported in their questionnaire responses.

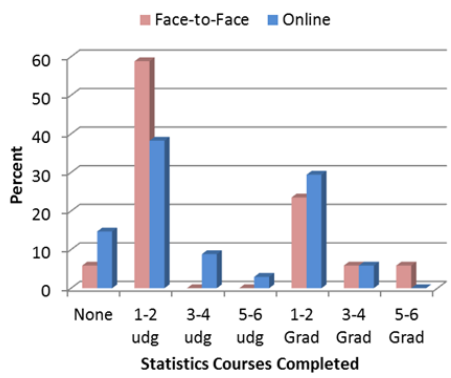
The online applicants were 57% European and 43% American. The face-to-face applicants were all Americans. Consequently we see a slightly higher self-reported English proficiency in the face-to-face participants. The face-to-face applicants were decidedly older than the online applicants. The face-to-face applicants had a slightly higher number of years of experience as a fire debris analyst. In all other areas, the groups were fairly well matched.

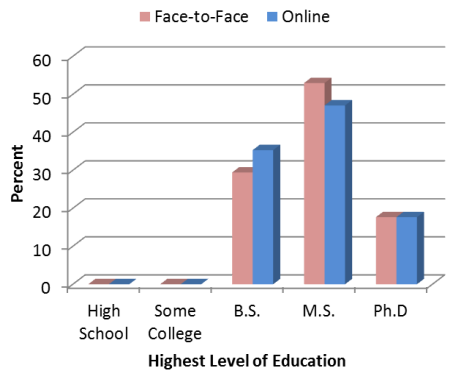


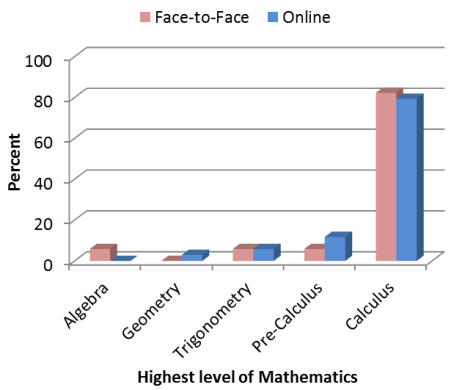


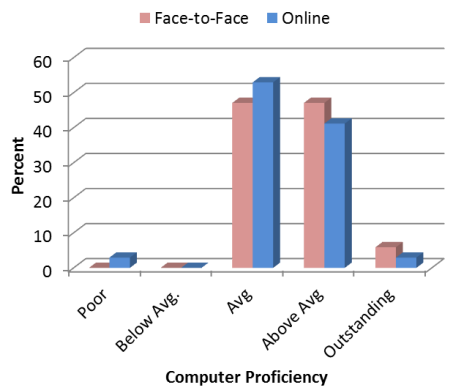
This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

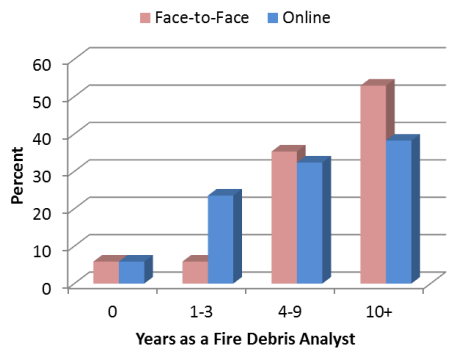












This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Appendix III: Face-to-Face Training Course Agenda

Training course in determining the probability for the presence of an ignitable liquid residue in fire debris by discriminant analysis	
Agenda	
8:00 AM – 8:30 AM	Introductions
8:30 AM – 9:15 AM	AMDIS Lesson Learn to download and install AMDIS, mass spectral software onto the computer. Learn to open a file, create a total ion spectrum and then save the total ion spectrum.
9:15 AM – 10:00 AM	R Lesson Learn to download and install R, a statistical software package onto the computer.
10:00 AM – 10:15 AM	Break
10:15 AM – 10:45 AM	TIS Lesson Learn about the origins of the total ion spectrum, appreciating the information content within the total ion spectrum, and recognizing theadvantages of utilizing the total ion spectrum for statistical data analysis methods.
10:45 AM – 11:30 AM	Discriminant Analysis Lesson Learn concepts of principal components analysis, linear discriminant analysis, and quadratic discriminant analysis.
11:30 AM – 1:00 PM	Lunch (on your own)
1:00 PM – 2:15 PM	Continue with Discriminant Analysis Lesson
2:15 PM – 3:00 PM	Putting It Altogether Lesson Apply the concepts learned to determine the probability for the presence of an ignitable liquid in fire debris.
3:00 PM – 3:15 PM	Break
3:15 PM – 4:00 PM	Continue with Putting it All Together Lesson
4:00 PM – 5:30 PM	Final Exam

This document is a research report submitted to the U.S. Department of Justice. This report has not been published by the Department. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.