



The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title: Substrate Pyrolysis Database Expansion, Evaluation, and Use

Author(s): Michael Sigman, Ph.D.

Document Number: 253892

Date Received: October 2019

Award Number: 2015-DN-BX-K051

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

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.

Agency: National Institute of Justice

Award number: 2015-DN-BX-K051

Project Title: Substrate Pyrolysis Database Expansion, Evaluation, and Use

PI: Michael Sigman, PhD
Director, National Center for Forensic Science
michael.sigman@ucf.edu
(407)-823-6469

Submitting official: Mary Davis
Award Management

Submission date: 10/11/2018

██████████ ██████████

██████████ ██████████

Recipient Organization: University of Central Florida
12201 Research Parkway Suite 501., Orlando, FL 32826

Award Period: 01/01/2016 to 12/31/2018

Final Summary Overview

Signature of Submitting Official: Mary Davis Digitally signed by Mary Davis
Date: 2018.10.11 08:20:54 -04'00'

Purpose

The purpose of the proposed developmental research was to produce an additional 1500-substrate pyrolysis records for the Substrate Database. This data assists fire debris analysts in casework, provides an extensive comparative cross tabulation of pyrolysis/combustion products to ignitable liquid components, and affords a more comprehensive understanding of the parameters that affect statistical classification models. Better statistical models result in lower error rates and improve likelihood estimates for the presence or absence of ignitable liquid residues (ILR) in fire debris samples. The goals of the research were to:

1. Develop a set of three standard burning methods that produce pyrolysis/combustion products.
2. Increase the number of records in the Substrate Database by preparing 10 records for 150 new substrate materials.
3. Conduct a cross tabulation of the frequency of occurrence based on identified compounds in ignitable liquids and pyrolysis/combustion products of substrate materials and a Naïve Bayes classification for the presence of ignitable liquid residue.
4. Develop better fire debris models to incorporate into statistical methods that allow analysts to assess objectively the evidentiary value of fire debris samples from casework.

Project design and methods

Described below is the process for generation and review of 1,500 new records into the Substrate Database. The SWGFEX Ignitable Liquids Database Collection (ILRC) committee reviewed the records during the course of the project for entry into the database. During the first meeting, committee members and the project group reviewed preliminary results, and established consensus burning methodologies for producing relevant data for the database. During the second meeting, committee members and the project group reviewed data, and created a best practices guide for the interpretation and use of database entries in casework and training.

Sample Preparation and Analysis of Substrate Materials

Four preliminary pyrolysis/combustion methods were developed and tested. The ILRC committee reviewed the results and provided recommendations for method modifications. These

recommendations led to the formation of three pyrolysis/combustion methods, which provided relevant data and information for the additional records entered into the Substrates Database. The tube furnace (fourth) method, based on previous work by Sandercock,[2] was eliminated because it produced similar results to the other methods, and is not described here. The following describes the final procedures for the modified destructive distillation (MDD), direct heat, and indirect heat methods. Fire debris forms under many influencing factors; these methods offer data that represent many of the possible results obtained in casework and span a significant portion of the data space in the statistical models. Substrate samples burned in this project consist of materials commonly recovered from households, commercial buildings, and automobiles collected at fire scenes.

Pyrolysis/Combustion Method 1 (MDDM)

The State of Florida Bureau of Forensic Fire and Explosives Analysis laboratory provided guidance for the development of the modified destructive distillation method (MDDM) based on their laboratory practices. Original Substrate database records utilized MDDM with a single burn-time interval of two minutes. The following describes the MDDM utilized for this project. Samples were created from substrate materials having dimensions of 16 cm² or weights of 1.5 grams. The sample was placed in an un-lined metal quart paint can and was positioned so the sample surface typically in contact with a heat source made contact with the bottom of the can, see Figure 1a. For example, the tufted or woven textile side of a carpet sample was placed in contact with the bottom of the can. Nine holes of approximately 1 cm in diameter were punched into a can lid, which was then placed on the can loosely. Heat produced from a propane torch was applied to the bottom of the can at a distance of 4 cm between the tip of the torch and the can to burn the sample. Heating was continued for an additional time interval after smoke appeared from the lid. After heating was discontinued, a solid lid replaced the punctured lid, allowing the vapors within the headspace to condense as the can returned to room temperature. A sample was prepared for each of three time intervals: 1, 2, or 5 minutes.

Pyrolysis/Combustion Method 2 (Direct Heat)

The direct heat method was based on a method developed by Sharee Booke Wells, MS, F-ABC (Forensic and Scientific Testing). The following describes the current Direct Heat method utilized for this project. The method involved placing the substrate sample onto the inner surface of a can lid positioned so the sample surface typically in contact with a heat source faced upward, see Figure 1b. A torch was used to ignite the sample while maintaining a distance of 4 cm and an angle of 75 degrees. Burning was continued for an interval of time, which began at contact of the flame to

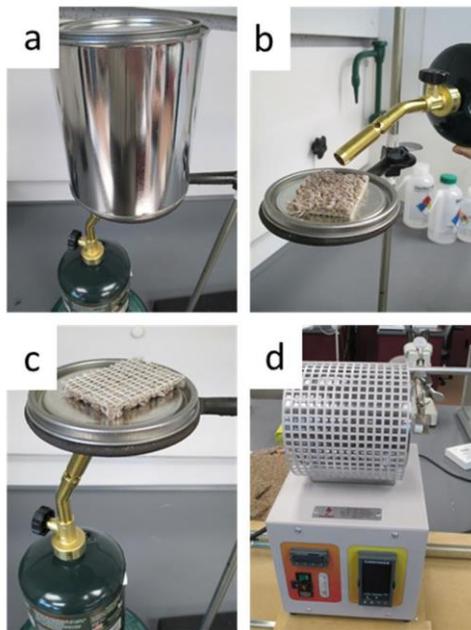


Figure 1. (a) Pyrolysis/combustion method 1, (b) method 2, (c) method 3, (d) method 4 (not described in the text).

the sample and then ceased upon smothering the burning sample. Smothering was achieved by placing a can over the sample and sealing it to the can lid. A sample was prepared for each of three time intervals: 1, 2, or 3 minutes.

Pyrolysis/Combustion Method 3 (Indirect Heat)

This method was a variation on Methods 1 and 2, discussed above. The following describes the Indirect Heat method utilized for this project. In this method, the pyrolysis/combustion of substrates proceeded by placing the materials on the inner surface of quart paint can lid and applying the heat from the bottom of the lid, see Figure 1c. Heating was continued for a period of time, which began once smoke appeared and ended with removal of the torch. Immediately afterwards, the sample and lid were covered with a can, and the sample was cooled to room temperature. A samples was prepared for each of three time intervals: 1, 2, or 3 minutes.

Extraction and Analysis

Pyrolysis and combustion products from the heated substrate materials were collected onto activated carbon by passive headspace adsorption following the ASTM E 1412 standard method. An activated carbon strip (1 cm x 3 cm) was suspended into the headspace of the can by a paperclip and un-waxed dental floss. The can was sealed and then heated for 16 - 18 hours at 66°C. After removal from the oven, the can was cooled to room temperature. Extraction of the adsorbed compounds was accomplished by submersion of the activated carbon strip in 0.5 ml of carbon disulfide. Analysis of the extract by gas chromatographic-mass spectrometry employed the method described online at the database website.

Compound Identification and Frequency of Occurrence for Major Peaks

The Ignitable Liquids Reference Collection (ILRC) database contains publicly available data for approximately 1,100 ignitable liquids collected over 18 years. The ILRC committee has identified five major (most intense) peaks for each liquid utilizing the NCFS mass-spectral library, which was created by analyzing standards using the same instrument method. Importantly, the sample and standard data were collected utilizing retention time locking, thus insuring that retention times were consistent. Major compound identification was verified by independent review. Major peak identification for samples in the Substrate Database followed the same process. The frequency of occurrence for each major compound in the library was calculated across 647 unevaporated ignitable liquids and 106 burned substrates (MDDM for 2 minutes).

Compound Identification-Frequency of Occurrence-Naïve Bayes Classification

Although the information gathered from the identification of major compounds is useful, it does not provide a comprehensive characterization of the ignitable liquids and substrate products. In this study, a greater number of compounds were identified utilizing the NCFS mass-spectral library of 295 standards. The National Institute for Standards and Technologies' (NIST) Automated Mass Spectral Deconvolution and Identification System (AMDIS) compared the mass spectra of unknown peaks in 643

unevaporated ignitable liquids and 522 substrates (burned by 1, 2, and 5 minute MDDM) with those in the library. Tentative peak identification was made for analytes with a spectral match factor of ≥ 80 . Verification was performed using logistic regression to obtain a probability of a matched based on retention time. The frequency of occurrence within ignitable liquids and substrates for each compound in the library was calculated. A Naïve Bayes classifier, which assumes variable independence, was used to calculate log likelihood ratios (LLR) for the presence on an ignitable liquid residue (ILR) from the frequencies of occurrence (probabilities) following equation 1.

$$LR = \frac{\prod_k P(C_k|H_{IL})}{\prod_k P(C_k|H_{SUB})} \quad \text{Equation 1}$$

Cross-validation (90% training and 10% testing) and performance metrics were employed to evaluate the performance of the Naïve Bayes classifier model using calibrated LLR. Validation consisted of training a new model with 1165 samples and testing with 16 fire debris samples of known ground truth (for the presence of ILR) that were similar and relevant to casework.

The Evidential Value of Fire Debris Data

Model-effects on likelihood ratios

The first study investigated model-effects on classification by distribution-directed random selection of 500 training samples from the ILRC and Substrate databases. Four principal component analysis (PCA) scores from the sample total ion spectra (TIS) were utilized as features to calculate log likelihood ratios by equation 2.

$$\log LR = \frac{|C_1|^{-1/2} \exp\left\{-\frac{1}{2}(\bar{y}-\bar{x}_1)^T (C_1)^{-1}(\bar{y}-\bar{x}_1)\right\}}{|C_2|^{-1/2} \exp\left\{-\frac{1}{2}(\bar{y}-\bar{x}_2)^T (C_2)^{-1}(\bar{y}-\bar{x}_2)\right\}} \quad \text{Equation 2}$$

Test data consisted of ten fire debris samples (MDDM for 0.5, 1, 2, 3, and 5 min.) of polyester carpet and carpet padding with one set of five samples containing gasoline and the other set containing a medium petroleum distillate (MPD).

Model distribution effects on likelihood ratios

The second study investigated what effects population and IL class distribution in a model have on the calculated likelihood ratios. The 10,000 training samples used for modeling consist of in-silico fire debris mixtures of TIS from 244 weathered IL, 445 unweathered IL, and 233 substrates (MDDM 2 min.).

Principal component analysis (PCA) scores that contain 90% of the variance in the training data became the factors used in Equation 1 to calculate log likelihood ratios assuming the covariance was the same (LDA) and different (QDA). The six (A – F) models developed consisted of different sample populations of IL and SUB as well as different IL distributions. Each model underwent a 10-fold cross-validation for both LDA and QDA.

Data analysis

Compound Identification and Frequency of Occurrence for Major Peaks

There were 185 unique major compounds identified at least once in ignitable liquids and substrates.

There were 102 compounds identified only in ignitable liquids, 47 identified only in substrates, and 36 identified in both. The major compounds in ignitable liquids consisted mostly of alkanes (normal, branched, cyclic, and bicyclic), aromatics, polynuclear aromatics, oxygenates (alcohols, ketones, and esters), and terpenes. Classes of compounds not identified in ignitable liquids were aldehydes, alkenes, imides, and styrenes. The major compounds identified in substrate pyrolysis consisted mostly of aldehydes, phenols, aromatics, alcohols, polynuclear aromatics, styrenes, esters, terpenes, alkenes, normal alkanes, and ketones. Classes of compounds not identified in substrates were cyclic alkanes, imides, and indanes. Thirty-one of the 36 compounds identified as major constituents in both ignitable liquids and substrates have different frequencies of occurrences. For example, n-none has a $2.49E^{-01}$ frequency of occurrence in ignitable liquids, but a frequency of $9.E^{-03}$ in substrates. Beta-pinene, 2-butoxyethanol, limonene, 2-methylnaphthalene, butyl acetate and have a similar frequency of occurrence in both ignitable liquids and substrates.

Compound Identification-Frequency of Occurrence-Naïve Bayes Classification

A Receiver Operating Characteristic (ROC) curve of the cross-validation results has an area under the curve (AUC) of 0.99, which approximates the method accuracy. An Empirical Cross Entropy (ECE) plot characterized the model as having good accuracy, calibration, and discriminating power. A Tippett plot and a Detection Error Trade-off (DET) plot demonstrated that the model had good discriminating power and the equal error rate is approximately 7%. The best accuracy involved employing only compounds identified in IL.

The Evidential Value of Fire Debris Data

Model-effects on likelihood ratios

Log likelihood ratios from cross-validation built ROC curves demonstrated the variability in the curves based on the randomly selected data for a model. Average AUC values were 0.959, 0.956, and 0.947 for each of the three distributions. Log likelihood ratios calculated for the fire debris testing data were positive for the 30-second burn time intervals and subsequently decreased to negative LLR for the 5-minute time interval that were statistically significant.

Model distribution effects on likelihood ratios

In cross-validation, distribution A for both LDA and QDA provided the most consistent AUC and better discrimination for all distributions (B-F). In addition, QDA had larger AUCs and better discrimination than LDA. Model F performed the worst because the training data set contained IL of only the gasoline class.

Findings

Substrate Database Modifications

Each substrate sample has ten related records in the database corresponding to the conditions of burning; one unburned sample, three burned by MDDM, three burned by direct heat, and three burned by indirect heat. Related records begin with an XX0 and end with XX9 where 0 – 2 are the MDDM, 3 – 5 are the direct heat, 6 – 8 are the indirect heat, and 9 is the unburned condition(s). Modification to the

database required two additional search fields pertaining to the method of burning and the time interval. In addition, the database needed to link the related records to each other for ease of searching and reviewing of the data.

Compound Identification and Frequency of Occurrence for Major Peaks

These frequencies of compound occurrence in ignitable liquids and substrates provides insight into the types of compounds that may be useful in characterizing fire debris. Identification of major compounds occurring only in ignitable liquids may assist in determining whether an ignitable liquid residue is present in a fire debris sample and the selection of target compounds for each ASTM E1618 class. Major compounds identified in both IL and SUB are not be considered indicative of the presence of an ignitable liquid residue in fire debris. This work was presented at the 2017 American Academy of Forensic Sciences and published in Forensic Chemistry July 2017 [1].

Compound Identification-Frequency of Occurrence-Naïve Bayes Classification

The Naïve Bayes classifier developed with identified compounds (i.e. chemical composition) of IL and SUB performed well during cross-validation utilizing data from IL and SUB. The method is well-calibrated, high discriminating power, and high accuracy; however, the accuracy is lower for the validation data set of fire debris samples. We recommend further pursuit of validation with a larger set of fire debris samples. This work was presented at the 2017 and 2018 American Academy of Forensic Sciences and the Australian and New Zealand Forensic Science Society's 2018 Symposia. A manuscript is in the process of being written.

The Evidential Value of Fire Debris Data

Model-effects on likelihood ratios

This study demonstrates the model-dependent variation in LLR values resulting from multiple random draws of training samples. The numerical objective method reports a mean LLR and confidence interval that reflects the strength of the fire debris evidence. Total ion chromatograms demonstrate a relationship between the ILR remaining in the sample and the strength of the evidence. This work was

published in the March 2018 edition of Forensic Chemistry and presented at Fall 2018 American Chemical Society National Meeting [2].

Model distribution effects on likelihood ratios

This study demonstrates that better performance of a model depends on the training data being representative of all sample populations and IL class distributions (especially uniform). A model cannot correctly predict the presence of an ILR if the ILR in the unknown sample is not represented in the training data used for development of the model. This work was presented at the 2018 American Academy of Forensic Sciences and was published in Separations special issue August 2018 [3].

Implications for Criminal Justice Policy and Practice in the United States

Expansion of the Substrates database is having a direct impact on fire debris casework in the United States and internationally by providing additional samples of pyrolysis products that were generated under different burn conditions. The NCFS fire debris databases (Ignitable Liquids Reference Collection, International Database of Ignitable Liquids and Substrates Database) are accessed on a regular basis by fire debris analysts around the world to assist in casework and to serve as training examples. The data from the expanded Substrate Database has improved computational models of fire debris and allowed for better predictions of likelihood ratios for the presence or absence of ignitable liquid residues in samples. The fire debris analysis community is resistant to change in data analysis protocols and practices, however, the improved modeling capabilities resulting from this work will facilitate this paradigm shift by the community and promote future use of probabilistic statements of evidential value, rather than the categorical statements used in current practice.

Works Cited

- [1] A. Akmeemana, M. R. Williams and M. E. Sigman, "Major chemical compounds in the Ignitable Liquids Reference Collection and Substrate databases," *Forensic Chemistry*, pp. 91-108, 2017.
<https://doi.org/10.1016/j.forc.2017.07.002>
- [2] R. Coulson, M. R. Williams, A. Allen, A. Akmeemana, L. Ni and M. E. Sigman, "Model-effects on likelihood ratios for fire debris analysis," *Forensic Chemistry*, vol. 7, pp. 38-46, 2018.
<https://doi.org/10.1016/j.forc.2017.12.008>
- [3] A. Allen, M. R. Williams, N. Thurn and M. E. Sigman, "Model distribution effects on likelihood ratios in fire debris analysis," *Separations*, vol. 5, no. 3, p. 44, 2018.
<https://doi.org/10.3390/separations5030044>