



Machine learning approaches over ion mobility spectra for the discrimination of ignitable liquids residues from interfering substrates[☆]

José Luis P. Calle^a, Barbara Falatová^b, María José Aliaño-González^a,
Marta Ferreiro-González^{a,*}, Miguel Palma^a

^a Department of Analytical Chemistry, Faculty of Sciences, Agrifood Campus of International Excellence (ceiA3), IVAGRO, University of Cadiz, Puerto Real, Cadiz 11510, Spain

^b Department of Fire Protection, Faculty of Wood Sciences and Technology, Technical University in Zvolen, ul. T. G. Masaryka 2117/24, Zvolen 960 53, Slovakia

ARTICLE INFO

Keywords:

Machine learning algorithms
Fire debris
Ignitable liquids
Ion mobility spectrometry
Petroleum-derived substrates

ABSTRACT

In arson fires, ignitable liquids (ILs) are frequently used to start combustion. For this reason, detecting IL residues (ILRs) at the fire scene is a key factor in fire investigation to determine whether a crime has been committed as well as to establish the modus operandi of the perpetrator. In the present study, the application of headspace-gas chromatography-ion mobility spectrometry (HS-GC-IMS) for the detection of ILRs in fire debris from complex matrices in combination with machine learning (ML) tools is proposed as an alternative to the traditional method, based on gas chromatography-mass spectrometry (GC-MS), described by the ASTM E1618 standard method. For this purpose, petroleum-derived substrates (vinyl, nylon, and polyester) and natural substrates (cotton, cork and linoleum) burned alone and with different ILs (gasoline, diesel, ethanol and charcoal starter with kerosene) were used. In addition, samples were taken at different times (0, 1, 6, 12, 24 and 48 h) after the fire was finished. The ion mobility sum spectrum (IMSS) of each sample was obtained and different ML algorithms were applied. The first derivative was performed at the IMSS, as well as a Savitzky-Golay filter. Hierarchical cluster analysis (HCA) revealed a clustering trend as a function of substrate and ILs used, where the studied sampling times did not affect the resulting clusters. The classification models for the detection of the presence of ILRs have high performance with an accuracy of 100% for support vector machines (SVM) and random forest model (RF), followed by linear discriminant analysis (LDA) with an accuracy of 86.67%. When discriminating the type of ILs used, the RF model obtained an accuracy of 100%, followed by the LDA with 97.22% and finally the SVM model with an accuracy of 93.06%. In addition, a simple web application has been developed where the trained models can be used, so any researcher can apply the method to detect ILRs in fire debris.

1. Introduction

Fire investigation is a discipline within Forensic Science and Criminalistics whose objective is to determine the origin and cause of a fire. When arson is suspected, the determination of the presence of a trace of ignitable liquids (ILs) or ignitable liquids residues (ILRs) at the fire scene is a key factor to establish the intentionality of the fire.

However, the detection of ILs at the fire scene, as well as the data interpretation, is still a challenging task that depends on numerous factors such as the destructive nature of the fire itself, the type of burned material, the type of ILs used, the firefighters' actions, the sampling time, or the weathering phenomena among others [1–3]. In recent years,

new ILs and household furnishing, especially those derived from petroleum, have appeared, resulting in more complex interferences since they share many signals with the ILs. Besides, the composition of the substrate may change during the combustion process due to pyrolysis of the compounds present in substrates, and also may be due to presence of ILRs. On the other hand, ILs might also result in different signals because the difference in calorific power causes different pyrolysis products [4]. All these factors highlight the difficulty of detecting and identifying the type of ILs used, especially when there are interfering substrates such as petroleum-derived products that may hinder the correct interpretation of data [5].

For these reasons, it is necessary to have robust analytical methods to

[☆] Edited by: Ian McKelvie

* Corresponding author.

E-mail address: marta.ferreiro@uca.es (M. Ferreiro-González).

detect the presence of ILs accurately. Up to now, the method par excellence is based on Gas Chromatography-Mass Spectrometry (GC-MS). In this sense, the American Society of Testing and Material (ASTM) has established a protocol for ILs analysis and classification into one of the eight classes established by GC-MS (E1618) [6]. This standard is usually coupled with a preconcentration method based commonly on the use of active carbon strips (ACS) established in ATSM - E1412 [7]. However, there are other methods of preconcentration that have been proposed such as solid phase microextraction [8], static headspace adsorption with tenax [9] or headspace sorptive extraction [10]. Regardless of the preconcentration method, the ASTM E1618 states that the identification of ILs is performed by visual comparison of the chromatogram, extracted ion profiling, and target compound with a reference database. Normally, each laboratory has its own library although they also use IL data collection libraries available online [11–13]. This methodology is somewhat subjective, time-consuming and requires a specialized operator. Therefore, more and more studies are proposing the use of chemometric tools to process the large amount of information obtained and automatically identify the presence of ILs with evidential strength (class belonging) [14]. Sigman et al. proposed the use of the TIS (total ion spectrum) where the identification of individual compounds in the sample is not carried out, and it has been successfully applied to determine the clustering tendency and to classify different ILs on different substrates, including petroleum-based products [15–18]. Moreover, it has been successfully applied even when different ILs altered by weathering and microbial degradation phenomena were used [19].

Most of the alternative methods to the ASTM E1618 are still based on GC-MS. In this sense, two-dimensional gas chromatography coupled to time-of-flight mass spectrometry have given excellent results [20] and has been tested on 19 different substrates with different ILs [21]. However, all these GC-MS methods also present some major disadvantages: long analysis times, higher cost, the use of contaminating solvents, sample preparation, or complexity of the technique among others. Therefore, other authors have evaluated several spectroscopic techniques as alternatives for the detection of ILs to reduce time, complexity and/or the use of solvents. In this case, González-Rodríguez et al. [4] and Kerr et al. [22] have obtained good results with Raman spectroscopy whereas Martín-Alberca et al. used Fourier Transform Infrared Spectroscopy [23] and Choi et al. applied laser-induced breakdown spectroscopy [24]. Ferreiro-González et al. have proposed the use of headspace-mass spectrometry electronic nose (HS-MS eNose), doing a direct analysis of the sample since there is no chromatographic separation nor identification of individual compounds [25]. With HS-MS eNose, the TIS is readily obtained and, in combination with the chemometrics used for the discrimination of samples, it is also obtained in a fast, eco-friendly and easy way. This methodology has been first successfully applied to thermally desorbed the ACS [26], and then for direct analysis of fire debris samples without the use of any adsorbent [25,27], including complex matrices samples [28] containing different fire suppression agents or that were subjected to weathering phenomena [29].

Recently, the potential of headspace analysis by headspace gas chromatography-ion mobility spectrometry (HS-GC-IMS) to detect the presence of ILs in fire debris has also been demonstrated [30,31]. Furthermore, it has been applied to study the biodegradation phenomena of different neat ILs [32]. This technique, in addition to being fast, cheap, eco-friendly and not requiring a specialized technician, has great portability because it works at atmospheric pressure, making it possible to perform *in-situ* analysis at the fire scene itself. However, this technique generates a large amount of data and, similar to TIS, the ion mobility sum spectrum (IMSS) is obtained. This has never been used for the identification of ILs in complex matrices such as burned petroleum-derived substrates or samples taken at different times, so the interpretation of the data requires advanced machine learning (ML) techniques that allow automating and simplifying the process.

Focusing on data analysis, the most commonly used chemometric

tools in most studies are principal component analysis (PCA), hierarchical cluster analysis (HCA) and linear discriminant analysis (LDA). However, the use of other techniques such as support vector machines (SVM) or random forest models (RF) is less commonly employed [33]. In addition, these techniques have never been evaluated in combination with IMSS to identify ILs. Nevertheless, a recent study has used SVM and RF models to determine the octane number in gasoline samples analyzed by HS-MS and NIR [34]. Furthermore, there are studies where the SVM model offers better results than the LDA in the identification of ILs in neat and ILs in fire debris, using the information corresponding to the TIS [18,35].

As it was previously mentioned, there is a Fire Debris Databases developed by the National Center for Forensic Science (NCFS) at the University of Central Florida (<https://ncfs.ucf.edu/databases>) [11–13] that contain characterization data used in the analysis of fire debris samples including neat ILs, ILRs and substrates, which have been analyzed following the ATSM E1618. These databases are very useful since they help forensic analysts to search for references samples to determine the presence and type of ILs in a test sample by visual comparison of the chromatogram. However, as previously discussed, the visual comparison established by ASTM 1618 has limitations. For this reason, many researchers are currently studying the possibility of applying mathematical algorithms (ML models) to assist the analyst in the interpretation of fire debris data [16]. So far, these models created by the researchers are not shared and therefore, other users are unable to automatically determine the presence of ILs in the sample. This makes the characterization of the samples challenging for the rest of the users, in addition to making interlaboratory validation difficult.

The objective of this study is to evaluate the suitability of different ML approaches (LDA, SVM and RF) applied on IMSS to automatically discriminate among different fire debris samples which contain interfering substrates (such as vinyl, nylon and polyester among others) that may hinder the correct interpretation of data [36]. Additionally, the samples were taken at different times in order to include more variability to the data. Gasoline, diesel, charcoal starter with kerosene and ethanol were used as IL to burn the materials.

2. Materials and methods

2.1. Samples: substrates and ignitable liquids

Six types of substrates were chosen for the fire debris preparation, three of them were petroleum-based products: vinyl flooring (43% ethylene and 57% chlorine), polyester carpet (100% polyester), nylon carpet (100% polyamide). The other three substrates, natural cork wood, 100% cotton sheets and linoleum flooring (limestone, wood powder, and linseed oil) were chosen for comparison purposes as they were natural. All of them are common to find in a residential home, so their analysis would be of great help to fire scene investigators and forensic scientists.

The ILs used for burning were purchased from local stores and these were: gasoline (95 Research Octane Number, lead-free), diesel (cetane index > 45), ethanol (absolute, 99.8% pure) from Panreac (Barcelona, Spain), and charcoal starter with kerosene (naphtha petroleum, hydro-treated heavy, kerosene). All the ILs were commonly found in arson fires due to their easy accessibility, availability and relatively low cost.

2.2. Fire debris sample preparation

The fire samples were obtained according to the protocol of the Destructive Distillation Method for Burning [37] with some modifications as it was previously reported [28]. Each substrate was burned in metal cans with each of the ILs described previously. Fire debris samples were taken from each can just after burning (10 min) and after 1, 6, 12, 24 and 48 h. The samples were kept in closed cans in a room with a constant temperature of 25 °C. Fire debris samples are taken at different

times because, in a real fire, sampling usually cannot be taken immediately, and it may require several days for the fire scene to be safe and examined. Therefore, the number of total samples was 360 (5 ILs/Alone x 6 substrates x 6 sampling times x 2 replicates).

The samples were labeled with the abbreviated name of the substrate used which was as follows: V (vinyl), LIN (linoleum), N (nylon carpet), PO (polyester carpet), CS (cotton sheet) and CO (cork). Subsequently, the code of the IL used for burning is added: Gas for gasoline, Die for diesel, Et for ethanol and Ker for charcoal starter with kerosene. Finally, the time of sample collection is indicated, and the label R1, R2 or P refers to the first and second replicate and an average of both, respectively. Thus, a cork sample burned without ILs and taken at 10 min would be labeled CO_OH_R1 for the first replicate, CO_OH_R2 for the second replicate and CO_OH_P to refer to the average of both replicates. In the case of a first replicate of a cotton sheet sample burned with gasoline and taken 48 h after the fire, it would be labeled as CS+Gas_48H_R1.

2.3. HS-GC-IMS analysis

All the fire debris samples were directly analyzed by using the FlavourSpec system (G.A.S., Dortmund, Germany). This unit consists of a headspace generator (HS) followed by a gas chromatography column (GC) coupled to an ion mobility spectrometry detector (IMS). The chromatographic column was very short being a multicapillary MCC OV-5 of 20 cm in length. Both the drift and carrier gas were nitrogen with a purity of 99.999%, and to create the headspace, a temperature of 80 °C and agitation for 20 min was applied. The rest of the analysis conditions for HS and GC-IMS can be found in the previously described literature for fire debris analysis [30]. Therefore, the total analysis time was 30 min.

2.4. Data analysis and software

Once all the samples were analyzed, the Ion Mobility Sum Spectrum (IMSS) was obtained using LAV HS-GC-IMS software (G.A.S., Dortmund, Germany). The IMSS is obtained by summing all the intensities for each drift time, independently of the chromatographic time. Therefore, the chromatographic information is eliminated. The range selected was from 1.029 to 1.92 relative to the Reaction Ion Peak (RIP) since this is the region where the characteristic signals of the samples appeared. The data was stored in 2D arrays represented by $D_{n \times p}$ where "n" refers to the number of samples and "p" to the number of variables (drift times). Thus, the final matrix will consist of 936 drift times (relative to RIP) and 360 samples, i.e., $D_{936 \times 360}$.

The analysis was performed using RStudio version 4.0.2 (RStudio Team 2021, Boston, MA, USA) but different packages were used, the main ones were: *prospectr* [38] to apply the derivative and the Savitzky-Golay filter, *factoextra* [39] for the hierarchical cluster analysis representations, *caret* [40] to train the different machine learning models and *shiny* [41] for application development.

The different machine learning models used are: I) Linear Discriminant Analysis (LDA) which is a supervised technique used for the classification of samples based on Fisher's linear discriminator. Discriminant functions are created by linear combinations of the original variables and are responsible for determining to which group each observation is assigned. II) random forest (RF) which is a supervised technique used in both regression and classification problems based on the generation of a multitude of individual trees. This algorithm includes hyperparameters that have to be previously chosen. These hyperparameters are the number of trees that make up the random forest model (known as *n_{tree}* hyperparameter) and the number of variables evaluated before splitting in each individual tree (known as *m_{try}* hyperparameter). III) Support Vector Machine (SVM) with Gaussian Kernel, which is a supervised technique used in both regression and classification problems based on the hyperplane concept. There are 2 hyperparameters whose value has to be previously chosen. One of them

is called cost (*C*) and controls the number of misclassifications allowed by the hyperplane, i.e., the number of support vectors. The other is called gamma (γ) and is responsible for controlling the flexibility of the model. The hyperparameter γ appears due to the use of a radial basis function (Gaussian kernel), which allows the generation of nonlinear separation limits.

3. Results and discussion

3.1. Exploratory analysis

All samples were analyzed by HS-GC-IMS, and spectra ranging from drift time 1.029 to 1.92 relative to the RIP were obtained. Throughout the exploratory analysis, the replicates were averaged to facilitate the visualization, therefore, the data matrix used is given by $D_{936 \times 180}$.

The raw IMSS for all the fire debris samples are represented in Fig. 1A. Since the objective is to detect first the presence or not of ILRs, as well as the type of ILR if there is any, the samples have been colored according to the type of IL used for burning. In this way, the IMSS of all samples belonging to each of the classes are shown: Substrates burned without ILs (Alone), substrates burned with gasoline (Gas_ILR), diesel (Die_ILR), ethanol (Et_ILR) and charcoal starter with kerosene (Ker_ILR).

As can be seen, the bands are wide and there is a baseline shift, especially large for the Ker_ILR samples. Focusing on the last region of the spectrum, for example, from drift time 1.769 to the end, it could easily distinguish the Ker_ILR samples from the rest since it presents a higher intensity. However, this is a consequence of the baseline shift due to the instrumental noise and slight variation in recording conditions and there may be no real differences between them. In order to develop robust predictive ML models and to avoid creating non-generalizable ILRs detection models, some data pretreatment was applied. The first derivative was calculated and a Savitzky-Golay smoothing filter (polynomial degree 3 and window size 11) was applied. The resulting spectra after such pretreatment are represented in Fig. 1B. It should be noted that after applying the Savitzky-Golay filter, the information edges are eliminated too and, therefore, the resulting matrix is then reduced from 936 to 882 variables ranging from drift time 1.034 to 1.915 (Relative to RIP).

As previously discussed, in Fig. 1A the Ker_ILR samples can be differentiated from the rest by observing the final region of the spectrum, but this is not possible in Fig. 1B as these are not real differences.

As can be seen, differentiating the samples visually based on their spectra is a difficult and subjective task. Besides, no clear differences are shown in all fire debris samples. In this case, the complexity of the matrix is too high and it is impossible to find clear generalizable patterns. It is noteworthy that some of the burned substrates are petroleum-derived, which gives an extra complexity because they could share some signals with the ILs used. In addition, it must be noted that samples with different post-fire periods of time were included in this data set, specifically from 10 min up to 48 h were used and during this time some chemical modification could happen.

For all these reasons, it is mandatory to use chemometric tools that help to obtain a better understanding of the structure and relationship of the data. In this case, a hierarchical cluster analysis (HCA) was carried out to assess whether there is any clustering tendency of the fire debris samples to be classified according to the presence /absence of the ILR and the type of ILR in case there is one. For this analysis, the matrix $D_{882 \times 180}$ corresponding to the average of the replicates and the use of the first derivative and Savitzky-Golay filter was used. In addition, the Ward's method with Manhattan distance was chosen. The resulting dendrogram (Fig. 2) is plotted in a circular format to ease its visualization.

To better understand the structure and relationships of the samples, in Fig. 2 the branches corresponding to the 8 main clusters have been colored and the two major clusters have been labeled with the letter "A" and "B". As can be seen, cluster A contains exclusively all samples from

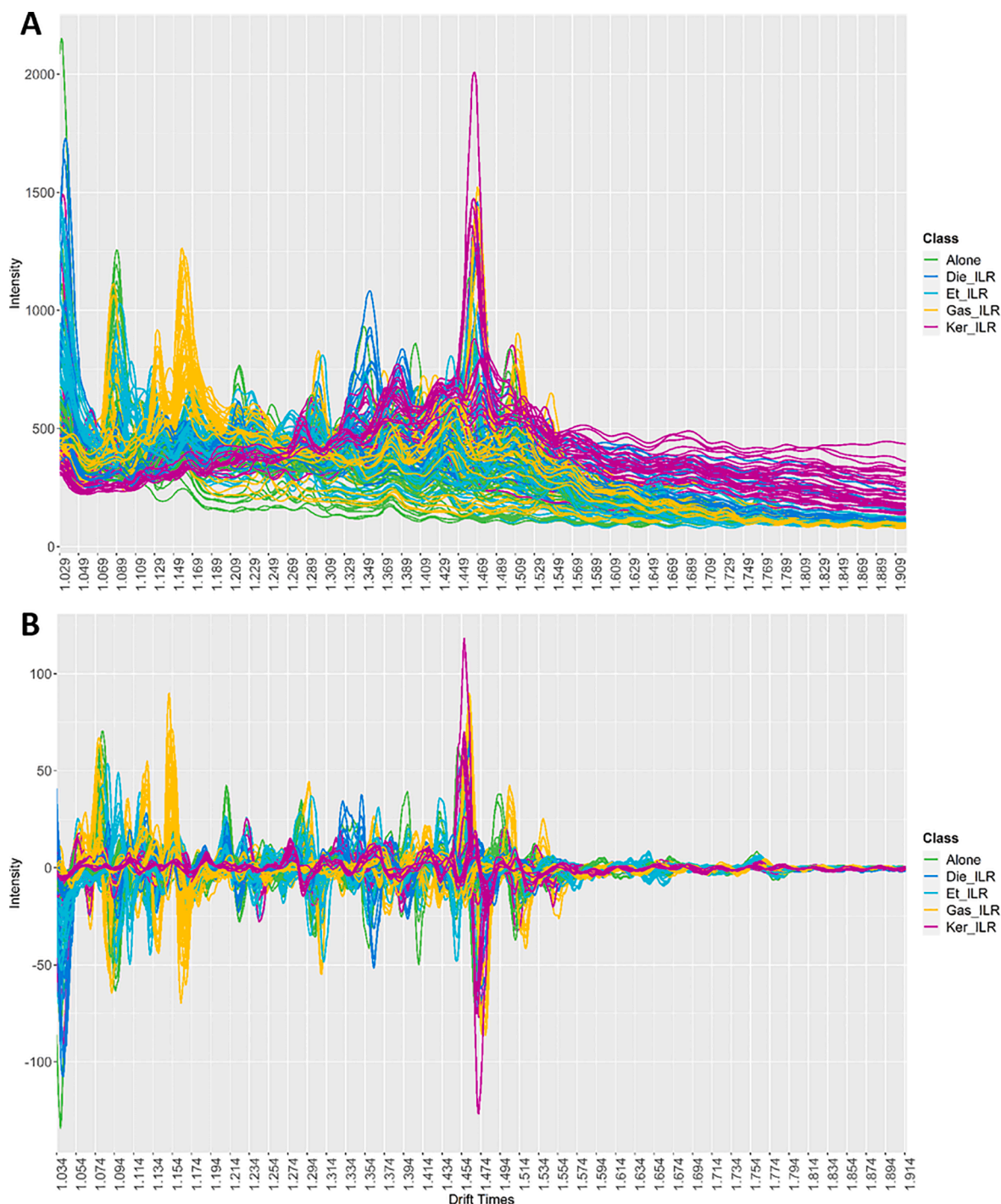


Fig. 1. (A) Representation of the raw spectra (IMSS) obtained for all the fire debris samples (D936x180) analyzed by HS-GC-IMS (B) IMSS for all the fire debris samples (D882x180) analyzed by HS-GC-IMS after the application of the first derivative and a Savitzky-Golay filter. All burned samples are colored according to the ILs used for burning: Alone for samples without IL (light green), Die_ILR for diesel (dark blue), Et_ILR for ethanol (light blue), Gas_ILR for gasoline (yellow) and Ker_ILR for kerosene (purple)

the alone group (without IL), as well as all samples from the Et_ILRs group. In the case of cluster B, there are all the samples burned with the rest of the ILs. Therefore, there is a strong tendency of classification according to whether the sample has been burned using ILs or not.

Focusing on cluster A, it can be seen that Et_ILR and Alone samples are mixed. Most of the ILs are complex mixtures of several compounds, however ethanol is a pure compound producing limited number of signals in the ion mobility spectra in the fire debris, regardless the substrate. Therefore, samples burned with ethanol are showing larger

similarities to samples burned without an IL than any other sample. Additionally, the substrate is very influential to the clustering results. The purple cluster contains the substrates LIN, V and PO. The brown cluster is made up of the N substrate samples, while dark green contains the CO and CS substrates. Therefore, the substrate itself as well as the combustion are the most important signals in the IMSS and are classified according to this.

Focusing on cluster B, a strong tendency to classify according to ILRs type as well as substrate is observed. So much so, that the dark blue

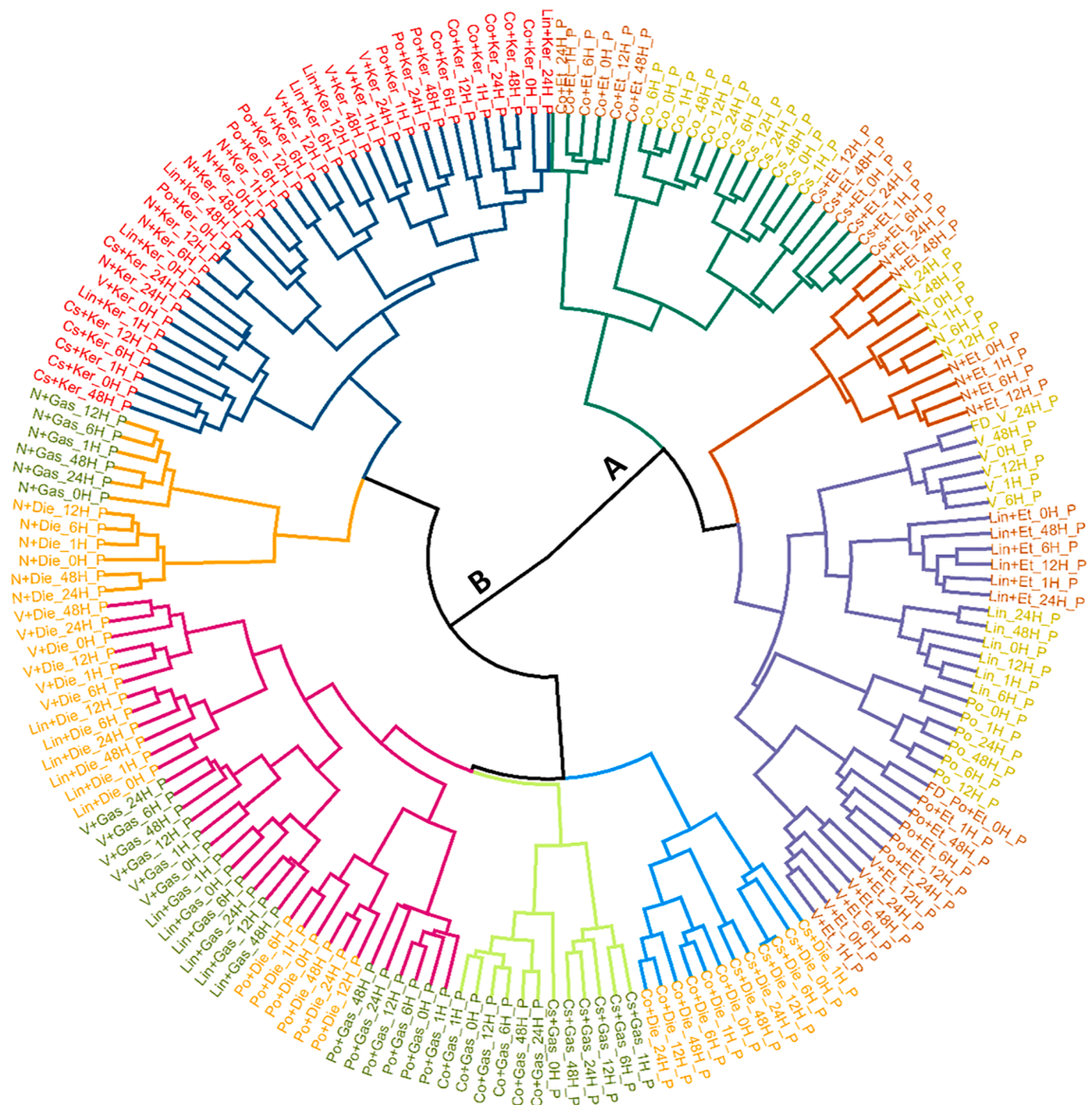


Fig. 2. Dendrogram resulting from the hierarchical cluster analysis (HCA) using Ward's method with Manhattan distances. Samples are colored according to the type of IL used for burning: Alone without IL (light green), Die for diesel (orange), Et for ethanol (brown), Gas for gasoline (dark green) and Ker for charcoal started with kerosene (red).

cluster is formed solely and exclusively by Ker_ILR samples. Furthermore, the different subclusters that are formed seem to differ mainly based on the substrate, for example, all Ker_ILR with CO substrate are in the same subgroup regardless of the sampling time. Subsequently, the group most similar to Ker_ILR is the orange cluster that contains the Nylon (N) samples burned with gasoline or diesel. In this case, the most important factor has been the substrate and the subcluster division is done according to the type of ILRs (gasoline or diesel). The sampling time does not seem to be relevant for this analysis. It should be noted that nylon is quite interfering when burned with gasoline [42], and at

high temperatures, volatile compounds characteristic of nylon are released. Therefore, the greater similarity of the ion mobility spectra of this group and the spectra obtained from the fire debris produced using kerosene, must be due to the chemical composition of nylon that is really different than the chemical composition of the other substrates.

It is observed that the violet cluster is formed by samples containing the LIN, V and PO substrates burned with both diesel and gasoline. However, the samples are divided into smaller groups depending on the type of ILs and substrate used, regardless of the sampling time. In addition, V and LIN substrates seem to be more similar to each other

than to PO, this could be due to the fact that vinyl and linoleum have a more similar chemical composition than polyester. Subsequently, the light green cluster contains the samples of the rest of the substrate (CO and CS) burned with gasoline, while the light blue cluster contains the same samples but burned with diesel. Once again, the trend of classification according to ILs and substrate is corroborated. Moreover, it is interesting that these two substrates (CO and CS) are separated from the rest (LIN, V and PO) as they are wood and cotton respectively, i.e., non-interfering and natural substrates. This is not the case of linoleum, which, despite being natural, presents a more complex ion mobility spectrum and so, it appears closer to those substrates derived from petroleum.

This data matrix was also analyzed by HS-MS and the result of the HCA analysis was similar [28]. In this case, the TIS was also influenced by the substrate and the type of ILRs while the sampling time did not seem to be as important. However, the subdivision as a function of substrate was not as remarkable as that obtained in IMSS.

These results suggest that the type of ILs for the ignition is an important factor that affects the IMSS and, in the case of use ethanol, there are no great differences with the samples burned without ILs, so in this case false negatives could be easily obtained. The substrate has a common signal with IMSS from ILs neat. In addition, clear differences between interfering substrates and natural ones (CO and CS) were observed. In the case of the studied sampling time, this does not seem to interfere with the classification of ILRs based on IMSS. This is a very useful aspect because sometimes in real fire scenes immediate sampling is usually not possible. In some cases, the actions may take several days, and this does not seem to be a problem when analyzed by the HS-GC-IMS technique. However, all these situations add extra complexity in identifying the presence of ILRs and the type used, but this complexity is necessary because samples need to reflect more accurately those taken in a real fire.

3.2. Supervised models for ILs detection

It has been shown that there is a tendency to classify according to the type of ILRs, but the variability due to the substrate burned makes it impossible to distinguish the presence of ILRs and its type perfectly. For this reason and, in order to have mathematical models to predict future samples, supervised models have been tested. Since in a real fire scene the first step is to determine whether an ILs is present or not, the models were created by establishing two classes a priori: "Alone" which includes the samples of the different substrates burned without ILs and "ILs" composed by the rest of the samples burned with the different ILs and substrates. Therefore, in this case, the "Alone" group is made up of 72 samples, which represents 20% of the total samples, while the "ILs" group has the remaining 288 samples, which represents the 80% of the total samples.

Different algorithms have been evaluated, including linear discriminant analysis (LDA), which is the most widely used technique in this type of research, as well as other non-parametric approaches based on support vector machines (SVM) with Gaussian kernel and random forest (RF). The full data matrix (with first derivative and Savitzky-Golay filter - D_{882x360}) was randomly split into two subsets with a ratio of 75% and 25% of the observations. In addition, it was ensured that class proportionality was kept in both subsets. The first one is the training set used to create the models and, the second one is the test set used to evaluate the performance of the models created. A summary of the accuracy obtained for the different trained models is shown in Table 1.

3.2.1. Linear discriminant analysis (LDA)

The LDA model was created with the training set, consisting of 75% of the total samples and using all the variables (882 drift times), as explained above. This LDA model was used to predict the samples used in its own creation (training set) and was able to correctly classify 100% of them. However, when the test set is predicted, 12 samples out of 90

Table 1

Summary of the accuracy obtained by the classification models tested for detection of ILRs.

Model	Hyperparameter	Training set accuracy	Test set accuracy
LDA	-	100%	86.67%
SVM	$C = 2.828$ $\gamma = 0.0098$	100%	100%
RF	$mtry = 30$ $ntree = 500$	100%	100%

were misclassified. This represents an accuracy of 86.67%, with a decrease in the performance of sample prediction never seen before by the model. In this case, the model suffers from overfitting, predicting the training set excellently but not being as generalizable to new observations. It should be noted that the groups are unbalanced and the "ILs" group constitutes the 80% of the data. Therefore, if the majority class is predicted, the accuracy will be 80% and the performance of this model should be considered quite low. In this case, it is better to use the kappa statistic to evaluate performance as it takes into account the probability that a prediction is correct simply by chance. The kappa value was 0.667, and according to the existing literature, values between 0.6 and 0.8 are considered substantial performance [43]. However, previous studies have obtained good performance of the LDA model in combination with IMSS when having less complex ILRs matrices [31]. Also, other studies using TIS from HS-MS analysis have reported good results when using the LDA model in complex matrices [18,28].

3.2.2. Random forest (RF)

The RF model was trained with the same training set as the LDA, but unlike this algorithm, the values for both hyperparameters explained above must be chosen beforehand. In this case, the $ntree$ was kept at 500 because it is a sufficiently large number to stabilize the error. The value of $mtry$ was 30 because for classification problems it is recommended to use the square root of the total number of variables [44].

Finally, the RF model was trained, leading to 100% of accuracy in the training set and 100% in the test set, therefore, the kappa statistic is 1. The 90 observations that constitute the test set were correctly classified and the probability of belonging to each class can be found in Table S1 (Supplementary material). It is observed that when the samples are burned without ILs the probability of correct classification is high, above 90% in most cases, and in samples burned with an IL the model always has a probability greater than 90% of identifying the presence of the ILs. Furthermore, the performance of the model is excellent (100% of samples classified correctly) and far superior to the LDA model.

3.2.3. Support Vector Machine (SVM) with Gaussian kernel

The SVM model was trained with the same training set as the previous models, and the hyperparameters C and γ were optimized. This optimization was performed in parallel and checked the accuracy by 5-folds cross-validation on the training set. Thus, models were created testing values of C and γ that grow exponentially, i.e., values every 0.5 in the range $[-10,10]$ were taken for $\log_2\gamma$ and \log_2C . This optimization approach has been used previously in numerous studies [45-47]. The accuracy obtained for the different combinations of hyperparameters is represented in 2D using a color map (Fig. 3). On the y-axis, the $\log_2\gamma$ is represented, while on the x-axis, the \log_2C and the accuracy are represented by a color scale where red represents the maximum value and blue the minimum.

As can be seen, the best results are obtained for low values of γ (approximately less than $\log_2\gamma = -6$, i.e., $\gamma = 0.0156$) and values of C greater than 0.5 ($\log_2C = -1$). In this case, the best combination was for $C = 4$ and $\gamma = 1.953 \cdot 10^{-3}$, leading to 100% accuracy in both the 5-fold cross-validation and the training set itself. In the case of the test set, the accuracy was 100%.

The probabilities of belonging for the test set samples in each class

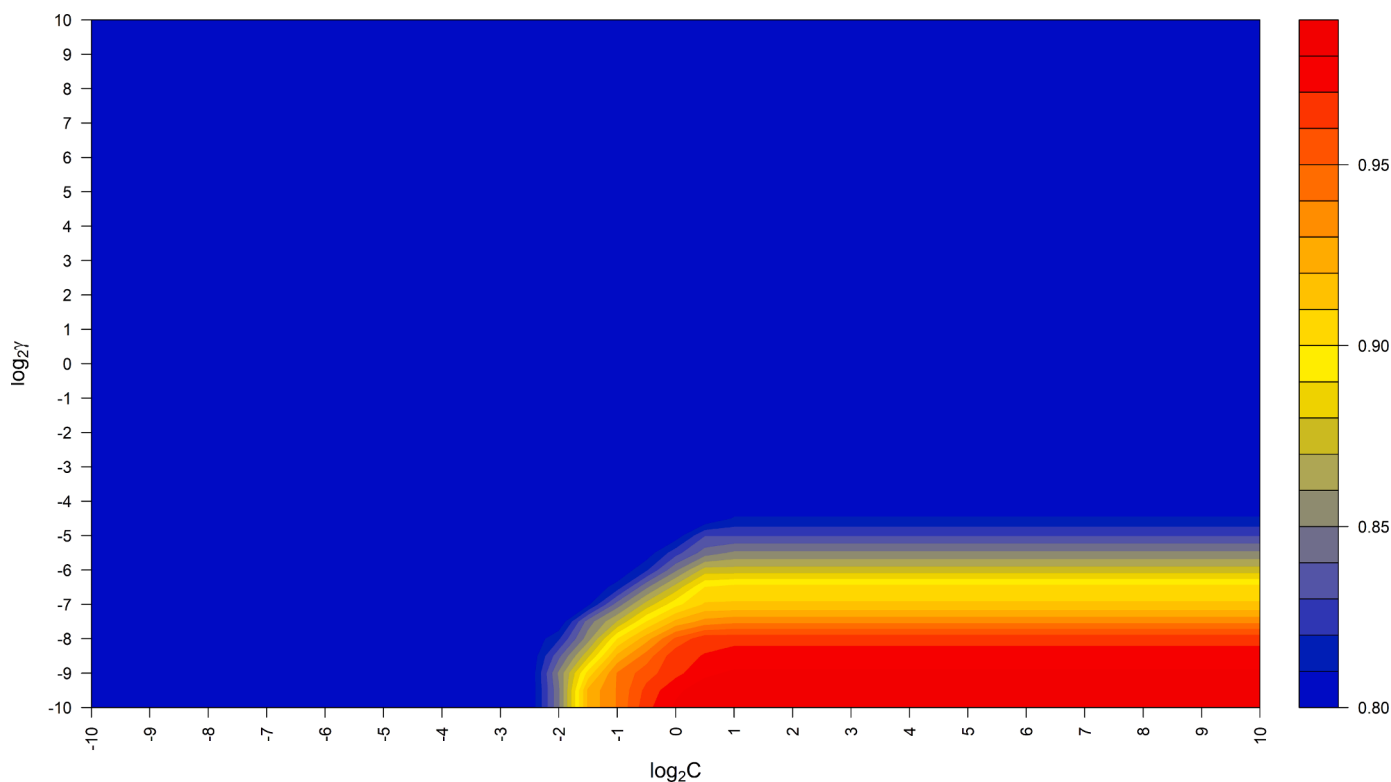


Fig. 3. Search for the best combination of hyperparameters (C and γ) for the Gaussian SVM model obtained by the CV of 5 folds using the IMSS of all the samples for the detection of ILs ($D_{882 \times 360}$).

are shown in Table S2 (supplementary material). Focusing on burned samples without ILs the probability of correct classification is in all cases higher than 93% (and mostly higher than 99%). Likewise, the probability of the correct classification of burned samples with ILs is very high, above 98% in all the samples.

In summary, both the RF and SVM models are excellent for detecting ILs, with 100% of accuracy for highly interfering samples such as petroleum-derived substrates. However, the LDA model is discarded, so both SVM and RF are appropriate to compare results. There are previous studies in the literature showing better results with the SVM or RF model than with the LDA model when analyzing fire residues using the information provided by the TIS [18,35].

3.3. Supervised models for ILRs identification

Once the models for detecting the presence of ILs in burned substrate have been created, the next step is to identify the type of ILs used. For this purpose, 4 classes must be established a priori corresponding to each type of ILs used: “gasoline”, “diesel”, “kerosene” and “ethanol”. In this case, each of the groups is made up of 72 samples, which represents a proportion of 25% of the total samples for each of the groups. The data matrix used is made up of all the ILRs samples (with first derivative and Savitzky-Golay filter - $D_{882 \times 288}$) which are divided into two subsets in the same way as described previously. The first subset is the training set consisting of 75% of the samples, and the second is the test set consisting of 25% of the remaining samples. Again, the models trained are the LDA, SVM with Gaussian kernel and RF. A summary of the accuracy is shown in Table 2.

3.3.1. Linear discriminant analysis (LDA)

The LDA model created with the training set obtained an accuracy of 100% in the training set and 97.22% in the test set. Therefore, 70 out of 72 observations were correctly classified. The 2 misclassified samples were as follows: Cs+Gas_24H_R1 predicted as diesel, Lin+Gas_6H_R1

Table 2

Summary of the accuracy obtained by the classification models tested for identification of different ILRs.

Model	Hyperparameter	Training set accuracy	Test set accuracy
LDA	-	100%	97.22%
SVM	$C = 1.414$ $\gamma = 1.381 \cdot 10^{-3}$	100%	93.06%
RF	$mtry = 30$ $ntree = 500$	100%	100%

predicted as diesel. The LDA model has obtained a good performance for the identification of the type of ILs used with a kappa statistic of 0.963. In addition, as can be seen the LDA model has only failed to predict the gasoline class as diesel. Furthermore, in Table S3 (supplementary material), other indices have been included, such as specificity, sensitivity, detection rate, detection prevalence and balance accuracy for each class.

3.3.2. Random forest (RF)

The RF model was created with the training set using the same hyperparameter values as in the previous model, i.e., a value of $ntree$ of 500 and $mtry$ of 30. This model obtained an accuracy of 100% for both the training set and the test set. Therefore, the value of the kappa statistic as well as the sensitivity, specificity and balance accuracy are 1. Table S4 (Supplementary material) shows the probabilities of belonging to each class. In all cases, these probabilities are high and generally greater than 80%. This indicates that the RF model obtains an excellent result for the identification of the type of ILRs.

3.3.3. Support vector machine (SVM)

Before training the SVM model, the hyperparameters were optimized in the same way as before. Fig. S1 (Supplementary material) shows the accuracy reached for each hyperparameter combination, obtaining the best value for a γ of $1.381 \cdot 10^{-3}$ ($\log_2 \gamma = -9.5$) and a C of 1.414 ($\log_2 C =$

-0.5). This combination of hyperparameters led to an accuracy of 100% in the training set and 93.06% in the test set. Therefore, the kappa statistic was 0.931 and other indices such as specificity, sensitivity, detection rate, detection prevalence and balance accuracy for each class can be found in Table S5 (supplementary material). In this case, there are 5 misclassified samples as follows: N+Die_0H_R1 which was predicted as kerosene, Cs+Ker_24H_R1 predicted as diesel, Lin+Ker_6H_R1 predicted as ethanol, Lin+Ker_12H_R1 predicted as ethanol and Lin+Ker_12H_R2 predicted as diesel. Misclassified samples burned with charcoal starter with kerosene predicted as diesel and vice versa are not so critical, since the ATSM - E1618 [5] considers these two ILs in the same class and even in the same subgroup, i.e., both are petroleum distilled and heavy (C₉-C₂₀+). In addition, Table S6 (supplementary material) shows the probabilities of belonging to each class. These probabilities are not as high as in the RF model. It is observed that although this model performs well, it suffers from some overfitting. Therefore, this model is not ideal for predicting the type of ILs used in fire debris analyzed by HS-GC-IMS.

3.4. Web application development

The models in combination with the IMSS have obtained excellent performance for the determination of the presence of ILs in fire debris as well as the type of ILs. For this reason, a simple application containing the SVM and RF models has been created to share the algorithm, as well as to facilitate the automatic identification of ILs for the rest of the users. This is very important since generally samples must be analyzed by GC-MS and the presence of ILs determined by visual comparison. In addition, the models are not usually shared so the use of the previously generated model would save time and cost, and can be accessed at the following link:

https://joseluispcalle.shinyapps.io/App_Fuegos_IMS/?_ga=2.47529639.1254510689.1641336025-1946327045.1641166303

This application is a prototype and automatically performs the first derivative and the Savitzky-Golay filter to subsequently make the prediction using the SVM model for the detection of ILs in fire debris. If the presence of ILs is detected, the prediction of the type (gasoline, diesel, ethanol and kerosene) is performed using the RF model. These two models have been used since they obtained an accuracy of 100%. Therefore, for its use, it is only necessary to upload the Excel file with the spectrum (IMSS) obtained from the analysis of the samples by HS-GC-IMS. In addition, an example file has been introduced that can be downloaded for free. The result is a table with the name of the sample and the prediction made by the models.

Although it is a very simple application, it allows to share the created algorithms and facilitate the analysis of fire debris samples to other users. It should be noted that both the application and the model can be improved with the analysis of more samples. In this way, the database will become progressively larger and will reflect better the scene of a real fire.

4. Conclusions

The potential of the HS-GC-IMS in combination with machine learning techniques for the detection and classification of ILs in fire debris in complex matrices has been demonstrated. The SVM and RF models have obtained the best results for the detection of ILs with an accuracy of 100% in the test set. However, in the identification of the type of ILs only the RF model obtained 100% of correct classification. In addition, the HCA analysis revealed that the substrate and type of ILs used greatly affect the IMSS, while sampling time does not affect it even two days after the fire. The HS-GC-IMS technique offers several advantages over traditional methods, such as short analysis time, low cost, simplicity and being environment friendly, as well as allowing *in-situ* analysis at the fire scene. An *in-situ* analysis is very important as it allows less manipulation of the sample and facilitates decision making at the

earliest possible time.

Funding

This work was co-financed by the 2014–2020 ERDF Operational Program; the Department of Economy, Knowledge, Business, and University of the Regional Government of Andalusia (FEDER-UCA18-107214); and 0756_FIREPOCTEP_6_E “Strengthening of cross-border systems for the prevention and extinction of forest fires and improvement of resources for the generation of rural employment post-COVID-19” (2021-020/PE/Interreg POCTEP 2014-2020/PR).

Declaration of Competing Interest

The authors declare that they have no conflict of interest.

Acknowledgments

José Luis Pérez Calle gratefully thanks the Ministry of Science and Innovation of Spain for a Ph.D. contract under the program FPU (FPU20/03377). The authors are grateful to the Instituto de Investigación Vitivinícola y Agroalimentario (IVAGRO) for providing the necessary facilities to carry out this research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.talo.2022.100125](https://doi.org/10.1016/j.talo.2022.100125).

References

- [1] J.R. Almirall, K.G. Furton, Characterization of background and pyrolysis products that may interfere with the forensic analysis of fire debris, *J. Anal. Appl. Pyrolysis* 71 (2004) 51–67, [https://doi.org/10.1016/S0165-2370\(03\)00098-6](https://doi.org/10.1016/S0165-2370(03)00098-6).
- [2] D.A. Turner, M. Williams, M.A. Sigman, J.V. Goodpaster, A Comprehensive study of the alteration of ignitable liquids by weathering and microbial degradation, *J. Forensic Sci.* 63 (2018) 58–65, <https://doi.org/10.1111/1556-4029.13527>.
- [3] J.S. Michael, K. Grau, M. Schmitz, J. Kim, Detection of gasoline residues present in household materials via headspace-solid phase microextraction and gas chromatography–mass spectrometry, *J. Anal. Chem.* 75 (2020) 44–55, <https://doi.org/10.1134/S1061934820010153/FIGURES/13>.
- [4] J. González-Rodríguez, N. Sissons, S. Robinson, Fire debris analysis by Raman spectroscopy and chemometrics, *J. Anal. Appl. Pyrolysis* 91 (2011) 210–218, <https://doi.org/10.1016/j.jaap.2011.02.012>.
- [5] J. Baerncopf, K. Hutches, A review of modern challenges in fire debris analysis, *Forensic Sci. Int.* 244 (2014) e12–e20, <https://doi.org/10.1016/J.FORSCIINT.2014.08.006>.
- [6] ASTM E1618-19, Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry, ASTM International, West Conshohocken, PA, USA, 2019.
- [7] ASTM E1412-19, Standard Practice for Separation of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Activated Charcoal, ASTM International, West Conshohocken, PA, USA, 2019.
- [8] ASTM E2154-15a, Standard Practice for Separation and Concentration of Ignitable Liquid Residues from Fire Debris Samples by Passive Headspace Concentration with Solid Phase Microextraction (SPME), ASTM International, West Conshohocken, PA, USA, 2015.
- [9] ASTM E1413-19, Standard Practice for Separation of Ignitable Liquid Residues from Fire Debris Samples by Dynamic Headspace Concentration onto an Adsorbent Tube, ASTM International, West Conshohocken, PA, USA, 2019.
- [10] J.I. Cacho, N. Campillo, M. Aliste, P. Viñas, M. Hernández-Córdoba, Headspace sorptive extraction for the detection of combustion accelerants in fire debris, *Forensic Sci. Int.* 238 (2014) 26–32, <https://doi.org/10.1016/J.FORSCIINT.2014.02.006>.
- [11] Fire Debris Database, National Center for Forensic Science, University of Central Florida, 2022 [internet][cited/03/27], Available from: <http://ilrc.ucf.edu/firedebris>.
- [12] Online Substrate Database, National Center for Forensic Science, University of Central Florida, 2022 [internet][cited/03/27], Available from: <http://ilrc.ucf.edu/substrate/>.
- [13] Ignitable Liquids Reference Collection Database, National Center for Forensic Science, University of Central Florida, 2022 [internet][cited/03/27], Available from: <http://ilrc.ucf.edu/>.
- [14] M.E. Sigman, M.R. Williams, Chemometric applications in fire debris analysis, *WIREs Forensic Sci* 2 (2020), <https://doi.org/10.1002/WFS2.1368>.

- [15] E.E. Waddell, E.T. Song, C.N. Rinke, M.R. Williams, M.E. Sigman, Progress toward the determination of correct classification rates in fire debris analysis, *J. Forensic Sci.* 58 (2013) 887–896, <https://doi.org/10.1111/1556-4029.12159>.
- [16] M. Lopatka, M.E. Sigman, M.J. Sjerps, M.R. Williams, G. Vivó-Truyols, Class-conditional feature modeling for ignitable liquid classification with substantial substrate contribution in fire debris analysis, *Forensic Sci. Int.* 252 (2015) 177–186, <https://doi.org/10.1016/J.FORSCIINT.2015.04.035>.
- [17] M.E. Sigman, M.R. Williams, J.A. Castelbuono, J.G. Colca, C.D. Clark, Ignitable liquid classification and identification using the summed-ion mass spectrum, *Instrum. Sci. Technol.* 36 (2008) 375–393, <https://doi.org/10.1080/10739140802151440>.
- [18] I. Sigman, M.W.- Assessing evidentiary value in fire debris analysis by chemometric and likelihood ratio approaches, *Forensic Sci.* 264 (2016) 113–121, <https://doi.org/10.1016/j.forsciint.2016.03.051>.
- [19] M.E. Sigman, M. Williams, *Degraded Ignitable Liquids Database: An Applied Study of Weathering and Bacterial Degradation on the Chromatographic Patterns of ASTM E 1618 Ignitable Liquid Classes*, National Center for Forensic Science, University of Central Florida, Orlando, 2016, 250468. NCJ Number.
- [20] K.D. Nizio, J.W. Cochran, S.L. Forbes, Achieving a near-theoretical maximum in peak capacity gain for the forensic analysis of ignitable liquids using GC×GC-TOFMS, *Separations* 3 (2016), <https://doi.org/10.3390/SEPARATIONS3030026>.
- [21] A.A.S. Sampat, B. Daelen, M. Van; Lopatka, H. Mol, G. Derweg, G.V. Van; Truyols, M. Sjerps, P.J. Schoenmakers, A.C.V. Astén, Detection and characterization of ignitable liquid residues in forensic fire debris samples by comprehensive two-dimensional gas chromatography, *Separations* 5 (2018), <https://doi.org/10.3390/SEPARATIONS5030043>.
- [22] T.J. Kerr, K.L. Duncan, L. Myers, Application of vibrational spectroscopy techniques for material identification from fire debris, *Vib. Spectrosc.* (2013) 225–235, <https://doi.org/10.1016/J.VIBSPEC.2013.08.006>. Complete.
- [23] C. Martín-Alberca, F.E.O. Ojeda, C. García-Ruiz, Study of spectral modifications in acidified ignitable liquids by attenuated total reflection fourier transform infrared spectroscopy, *Appl. Spectrosc.* 70 (2016) 520–530, <https://doi.org/10.1177/0003702815626681>.
- [24] S. Choi, J.J. Yoh, Fire debris analysis for forensic fire investigation using laser induced breakdown spectroscopy (LIBS), *Spectrochim. Acta Part B At. Spectrosc.* 134 (2017) 75–80, <https://doi.org/10.1016/J.SAB.2017.06.010>.
- [25] M. Ferreiro-González, G.F. Barbero, M. Palma, J. Ayuso, J.A. Álvarez, C.G. Barroso, Determination of ignitable liquids in fire debris: direct analysis by electronic nose, *Sensors* 16 (2016), <https://doi.org/10.3390/S16050695> (Basel).
- [26] M. Ferreiro-González, J. Ayuso, J.A. Álvarez, M. Palma, C.G. Barroso, Application of an HS-MS for the detection of ignitable liquids from fire debris, *Talanta* 142 (2015) 150–156, <https://doi.org/10.1016/J.TALANTA.2015.04.030>.
- [27] M. Ferreiro-González, G.F. Barbero, J. Ayuso, J.A. Álvarez, M. Palma, C.G. Barroso, Validation of an HS-MS method for direct determination and classification of ignitable liquids, *Microchem. J.* 132 (2017) 358–364, <https://doi.org/10.1016/J.MICROC.2017.02.022>.
- [28] B. Falatová, M. Ferreiro-González, J.L.P. Calle, J.Á. Álvarez, M. Palma, Discrimination of ignitable liquid residues in burned petroleum-derived substrates by using HS-MS eNose and chemometrics, *Sensors* 21 (2021) 1–12, <https://doi.org/10.3390/s21030801>.
- [29] B. Falatová, M. Ferreiro-González, C. Martín-Alberca, D. Kačíková, Š. Galla, M. Palma, C.G. Barroso, Effects of Fire Suppression Agents and Weathering in the Analysis of Fire Debris by HS-MS eNose, *Sensors* 18 (2018), <https://doi.org/10.3390/S18061933> (Basel).
- [30] M.J.A. González, M.F. González, G.F. Barbero, M. Palma, C.G. Barroso, Application of headspace gas chromatography-ion mobility spectrometry for the determination of ignitable liquids from fire debris, *Separations* 5 (2018) 41, <https://doi.org/10.3390/SEPARATIONS5030041>, 2018, Vol. 5, Page 41.
- [31] M.J. Aliño-González, M. Ferreiro-González, G.F. Barbero, M. Palma, Novel method based on ion mobility spectrometry sum spectrum for the characterization of ignitable liquids in fire debris, *Talanta* 199 (2019) 189–194, <https://doi.org/10.1016/J.TALANTA.2019.02.063>.
- [32] J.L.P. Calle, M. Ferreiro-González, M.J. Aliño-González, G.F. Barbero, M. Palma, Characterization of biodegraded ignitable liquids by headspace-ion mobility spectrometry, *Sensors* 20 (2020) 6005, <https://doi.org/10.3390/S20216005>.
- [33] M.W. ME Sigman, *Chemometric Applications in Fire Debris Analysis*, 2, Wiley Online Libr, 2020, <https://doi.org/10.1002/wfs2.1368>.
- [34] M. Barea-Septúlveda, M. Ferreiro-González, J.L.P. Calle, G.F. Barbero, J. Ayuso, M. Palma, Comparison of different processing approaches by SVM and RF on HS-MS eNose and NIR Spectrometry data for the discrimination of gasoline samples, *Microchem. J.* (2022) 172, <https://doi.org/10.1016/J.MICROC.2021.106893>.
- [35] M.E. Sigman, M.R. Williams, N. Thurn, T. Wood, Validation of ground truth fire debris classification by supervised machine learning, *Forensic Chem.* 26 (2021), 100358, <https://doi.org/10.1016/J.FORC.2021.100358>.
- [36] P.M.L. Sandercock, Background Interference in Fire Debris Analysis, in: K. Evans-Nguyen, K. Hutches (Eds.), *Forensic Analysis of Fire Debris and Explosives*, Springer, Cham: Switzerland, 2019, https://doi.org/10.1007/978-3-030-25834-4_3.
- [37] M.R. Williams, M.E. Sigman, J. Lewis, K.M.H. Pitan, Combined target factor analysis and Bayesian soft-classification of interference-contaminated samples: forensic fire debris analysis, *Forensic Sci. Int.* 222 (2012) 373–386, <https://doi.org/10.1016/J.FORSCIINT.2012.07.021>.
- [38] Stevens A, Ramirez-Lopez L; Prospecr: An introduction to the prospecr package. 2022; Available online: <https://cran.r-project.org/web/packages/prospecr/index.html> (accessed on 18 January 2022).
- [39] Kassambara, A. and Mundt, F; Factoextra: extract and visualize the results of multivariate data analyses. 2020. Available online: <https://CRAN.R-project.org/package=factoextra> (accessed on 18 January 2022).
- [40] Kuhn, M; caret: classification and regression training. 2020; Available online: <https://cran.r-project.org/web/packages/caret/caret.pdf> (accessed on 18 January 2022).
- [41] Chang, W.; Cheng, J.; Allaire, J.J.; Xie, Y.; McPherson, J; Shiny: web application framework for R. 2020; Available online: <https://cran.r-project.org/web/packages/shiny/index.html> (accessed on 19 February 2022).
- [42] E.M. Hondrogiannis, C. Newton, R. Alibozek, Determining the method threshold of identification via gas chromatography–mass spectrometry of weathered gasoline extracted from burnt nylon carpet, *J. Forensic Sci.* 64 (2019) 1160–1168, <https://doi.org/10.1111/1556-4029.13983>.
- [43] B. Lantz, *Machine Learning with R*, 2nd ed., Packt Publishing Ltd, 2015 [Book] ISBN 9781784393908.
- [44] Aurélien Géron, *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, 2nd ed., O'Reilly Media, Inc., 2019 [Book] ISBN 9781492032649.
- [45] J.L.P. Calle, M. Ferreiro-González, A. Ruiz-Rodríguez, G.F. Barbero, J.Á. Álvarez, M. Palma, J. Ayuso, A methodology based on FT-IR data combined with random forest model to generate spectralprints for the characterization of high-quality vinegars, *Foods* 10 (2021) 1411, <https://doi.org/10.3390/foods10061411>.
- [46] S. Qiu, J. Wang, L. Gao, Discrimination and characterization of strawberry juice based on electronic nose and tongue: comparison of different juice processing approaches by LDA, PLSR, RF, and SVM, *J. Agric. Food Chem.* 62 (2014) 6426–6434, <https://doi.org/10.1021/jf501468b>.
- [47] H. Men, S. Fu, J. Yang, M. Cheng, Y. Shi, J. Liu, Comparison of SVM, RF and ELM on an electronic nose for the intelligent evaluation of paraffin samples, *Sensors* 18 (2018) 285, <https://doi.org/10.3390/S18010285>.