Study of the influence of the matrix on the source inference process of ignitable liquids by an untargeted chemometric approach

Insaf BEN TERDAYET[1]   Xavier ARCHER[1]   Jean-marc BEGUE[1]   Olivier DELEMONT[2]   Anne Varenne[3]

[1] Laboratoire Central de la Préfecture de Police, 39bis rue de Dantzig 75015 Paris France, insaf.ben-terdayet@interieur.gouv.fr, Xavier.archer@interieur.gouv.fr, Jean-marc.Begue@interieur.gouv.fr

[2] Ecole des Sciences criminelles, Université de Lausanne, CH-1015 Lasanne, Suisse, Olivier.Delemont@unil.ch

[3] Chimie ParisTech PSL, 75005 Paris, France, Anne.varenne@chimieparistech.psl.eu

Keywords: Source inference, ignitable liquids, ROC curve, GC-MS

1   Introduction

Identifying the perpetrator of arson is a real challenge for police investigators and forensic scientists since the physical traces abandoned by the perpetrator on this type of crime scene are often destroyed. If an ignitable liquid was used to start the fire, and its residual presence can be detected, a source inference approach with other ignitable liquids is likely to provide investigative or evidentiary support. Source inference is a process of association that consists of evaluating the possible common source between two ignitable liquids specimens, in particular between a liquid extracted from a debris collected at the arson site and a liquid found in an external source (jerry can, clothes found in the vicinity...). This inference is mainly based on the comparison of similarities and/or differences in the chemical composition of the ignitable liquid specimens being investigated. This comparison is easier when the ignitable liquid specimens are in their liquid form, pure or even partially altered. However, due to the destructive nature of fires, the samples received in the laboratory for analysis consist mainly of solid materials in which liquid residues have sorbed, and most often these samples have been physically and chemically altered by the fire.Previous research has shown that it was possible to link two altered or unaltered ignitable liquids sharing a common source using an untargeted chemometric approach on GC-MS results[1]. Yet the challenge of considering solid fire debris containing ignitable liquid residues add a significant level of difficulty. The matrix effect, which includes the effect of any substrate present in the debris, is extremely important in this type of specimen, since compounds from the substrates can compromise the extraction of ignitable liquids, produce numerous interfering compounds from the raw material, its combustion or its pyrolysis, affecting the interpretation of the presence or absence of ignitable liquids in the analyzed debris. As a result, an alteration of the profiles is systematically observed, which increases the level of complexity of a source inference approach. In previous works, the analysis of liquid gasoline samples allowed to link two liquid samples independently of their mode and degree of alteration thanks to the development of a similarity calculation algorithm[1]. In this work, we propose to study the matrix contribution to the problem of source inference in forensic science by using the multivariate chemometric methods already developed.

2   Material and methods

First, the selection of gasoline samples (representative of very different chemical compositions) was carried out by a first screening by principal component analysis (PCA). The choice of the matrix was focused on clothing (100% cotton or 100% polyester T-shirts), little studied until now, but often found as samples in fire debris analysis. Simulated fire debris specimens were produced in

laboratory conditions with these gasoline and fabrics samples. They were then analyzed by thermodesorption followed by gas chromatography coupled to single quadrupole mass spectrometry (TD-GC-MS, in full scan mode). A data processing and analysis step is then performed: after a data pre-processing phase (detection, integration and alignment of chromatographic peaks), the variables of interest (areas, retention time and m/z) are then extracted in the form of data matrix which are processed by an algorithm under Matlab (R2019b). This algorithm builds a model based on the ROC (Receiver operating characteristics) curve using a large dataset of gasoline samples from known sources, taking into account the intra- and inter-variability of these samples. This model may then be applied to unknown gasoline samples to assess a possible link.

3   Results and discussion

An example of a model applied on gasoline samples extracted from cotton is shown below:
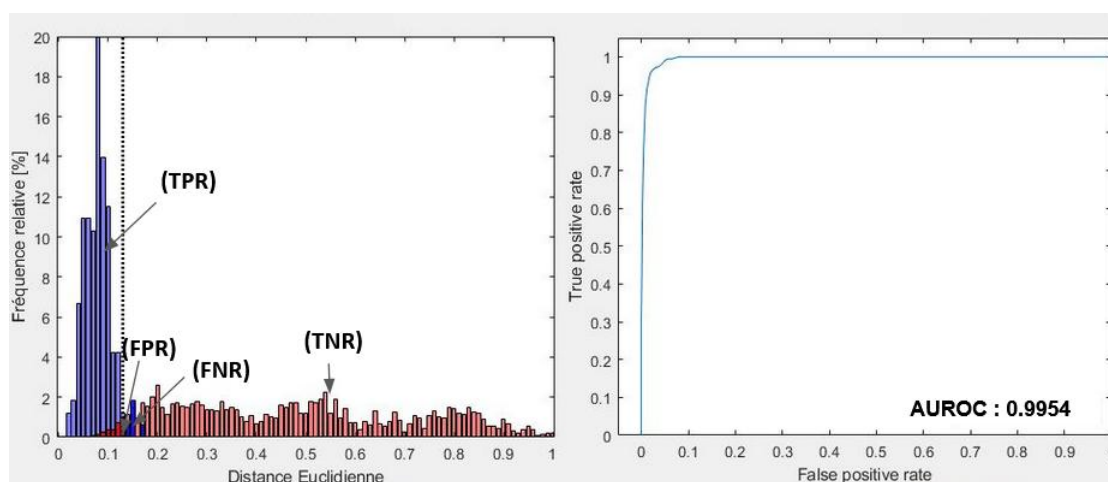
Figure 1: Final distribution (model) of linked and unlinked gasoline samples extracted from cotton fabric as a function of Euclidean distance (*Left*), ROC curve (*Right*)

The algorithm classifies correctly the gasoline samples sharing a common source extracted from cotton (Very low false positive (FPR) and false negative rates (FNR)). The area under the ROC curve (AUC) indicates the performance of the constructed model[2]. As described above, the value of the AUC is very close to 1 which confirms the relevance of the result obtained.

4   Conclusion

After identifying the signals from the matrix, we have shown that it is possible to link two gasoline samples altered by evaporation extracted from the fabric (Cotton and polyester), working on a small number of samples. These results allowed us to extend the study path for more complex cases, studying the effect of burned gasoline samples in the presence of a burned matrix on a source inference approach.

5   References

[1]   De Figueiredo M, Jouan-Rimbaud Bouveresse D, Cordella CBY et al : *Exploratory study on the possibility to link gasoline samples sharing a common source after alteration by evaporation or combustion*. Forensic Science International: 301:190-201, 2019

[2]   J A Hanley, B J McNeil : *The meaning and use of the area under a receiver operating characteristic (ROC) curve.* Radiology, 143: 29-36, 1982