

Multivariate Statistical Analysis in Fire Debris Analysis

Viacrozmerné štatistické metódy pri analýze vzoriek z požiariska

Barbara Falatová^{1*}, Marta Ferreiro-González², Danica Kačíková¹, Štefan Galla³, Miguel Palma², Carmelo G. Barosso²

¹ Department of Fire Protection, Faculty of Wood Sciences and Technology, Technical university in Zvolen, ul.T. G. Masaryka 2117/24, 960 53 Zvolen, Slovak Republic; barbara.falatova@gmail.com; danica.kacikova@tuzvo.sk;

² Department of Analytical Chemistry, Faculty of Sciences, University of Cadiz, Agrifood Campus of International Excellence (ceiA3), IVAGRO, P.O. Box 40, 11510 Puerto Real, Cadiz, Spain; marta.ferreiro@uca.es; miguel.palma@uca.es; carmelo.garcia@uca.es;

³ Fire Research Institute of the Ministry of Interior, Rožňavská 11, 831 04 Bratislava, Slovak Republic; stefan.galla@minv.sk;

* Corresponding author: barbara.falatova@gmail.com

Short Communication

Received: June 18, 2018; Accepted: July 7, 2018; Published: July 31, 2018;

Abstract

An U. S. national study demonstrated an increased reliance on statistical methods to data analysis in Forensic Sciences. The multivariate statistical analysis is recommended, in order to simplify the processing and understanding the data obtained from analytical techniques. In recent years, there has been an increasing interest in applying chemometric tools for data interpretation in fire investigation. The paper aims to present a progressive approach to data analysis by a brief description of multivariate statistical methods that are widely used for data analysis and data interpretation of fire debris. The paper describes chemometric approach to data analysis and gives examples of successful application of chemometrics to fire debris analysis.

Keywords: Multivariate statistical analysis; Chemometrics; Fire Debris Analysis

1 Introduction

Fire investigation also referred to as origin and cause determination is a forensic science that covers knowledge from many other disciplines. Preliminary scene examination, evidence collection and laboratory analysis, hypothesis evaluation and conclusion belong to basic steps of the fire investigation process [1].

The laboratory analysis is an important step in the process of fire investigation and plays a key role in determining the purpose of a fire. The main challenge faced by many forensic chemists is data interpretation from fire debris samples. The data interpretation plays an important role in addressing the role of the

1 Úvod

Zisťovanie príčin vzniku požiarov (ZPP) je multidisciplinárna veda, ktorá zahŕňa v sebe vedomosti z viacerých oblastí. Medzi základné kroky v procese ZPP patrí: obhliadka požiariska, odoberanie vzoriek a laboratórna analýza, tvorba a hodnotenie hypotéz a tvorba záveru [1].

Vzorky odobrané z požiariska sú počas ZPP podrobené laboratórnej analýze. Zisťuje sa napríklad, prítomnosť zvyškov urýchľovačov horenia. Pre relevantnosť výsledkov, vydala medzinárodná organizácia pre rozvoj a normalizáciu (ASTM) normy a štandardy týkajúce sa metodiky prípravy, extrakcie a analýzy vzoriek z požiariska.

possible use of an accelerant. Gas chromatography-mass spectrometry (GC-MS) is the analytical method, most widely used around the globe [2, 3].

Studies over the past two decades have provided important information on distortion effects and interfering compounds that may hinder or entangle the interpretation of chromatographic data [2]. However, a major problem with data analysis is, that it is subjected to human interpretation, caused by visual comparison of obtained chromatograph with the reference database [4].

There are three methods of data analysis for ignitable liquid (IL) classification – chromatographic pattern recognition (by using total ion chromatogram), extracted ion profiling, and target compound analysis.

One of the strains of fire debris classification from different laboratories may be caused by comparison of chromatographic data, specifically total ion chromatograms (TICs) [5]. On the other hand, total ion spectrum (TIS) provides an alternative approach for data analysis. The previous study proved [5], that the TIS is identical to an average mass spectrum covering the complete chromatographic range.

The aim of the study is to describe multivariate statistical analysis and their application on fire debris.

2 Novelty in fire debris analysis

The concept of interpretation has been challenged by U. S. national study [6] demonstrating an increased reliance on statistical methods to data analysis among forensic sciences. The multivariate statistical analysis is recommended, in order to simplify the processing and understanding of data obtained from analytical techniques.

In recent years, there has been an increasing interest in applying chemometric tools for data interpretation in fire investigation. Previous research has established that chemometrics is a useful tool in data analysis and may successfully determine the presence of ignitable liquid residues [2].

Nevertheless, the most of forensic laboratories perform fire debris analysis by GC-MS, several authors have been working on alternative methodologies that can complement or may become an alternative to traditional methods because of advantages they provide.

Plynová chromatografia s hmotnostnou spektrometriou (GC-MS) je analytická metóda, ktorá sa vo forenzných laboratóriách využíva najčastejšie [2,3]. V priebehu posledných rokov došlo k niekoľkým zlepšeniam v súvislosti s touto metodikou. Boli navrhnuté alternatívne prístupy, aby sa prekonali niektoré nevýhody, ktoré táto metodika predstavuje [2]. Podľa Borusievitza [4] je táto metodika podrobená prevažne subjektívnej interpretácii, pretože je založená na vyhodnotení celkového iónového chromatogramu (TIC) alebo extrahovaného iónového chromatogramu (EIC) hlavných zlúčenín vizuálnym porovnaním vzoriek s referenčnou databázou [4]. Okrem toho, táto metodika je časovo náročná a interpretácia výsledkov sa stáva zložitejšou, pretože vzorky nie sú čisté kvapaliny, ale obsahujú veľké množstvo materiálov.

Problémy súvisiace s interpretáciou dát sú napríklad: zvetrávanie, mikrobiologická aktivita alebo interferujúce zlúčeniny. Okrem toho je komplikované klasifikovať vzorky aj kvôli prítomnosti iných prchavých zlúčenín pochádzajúcich zo substrátu, produktov spaľovania alebo pyrolýzy [2].

Cieľom príspevku je popis viacrozmerných štatistických metód a ich aplikácia na analýzu vzoriek z požiariska.

2 Progres v analýze vzoriek z požiariska

Podľa americkej národnej štúdie [6], sa odporúča a kladie sa dôraz na použitie štatistických metód vo forenzných vedách. Viacrozmerné štatistické analýzy môžu zjednodušiť spracovanie a interpretáciu dát získaných z analytických techník [2]. Použitie chemometrických nástrojov analýzy sa v posledných rokoch zvýšilo.

Waddell *et al.* [5] uviedli, že jeden z aspektov sťaženej detekcie akceleratorov môže byť spôsobený porovnaním údajov z chromatogramu, konkrétne TIC. Na druhej strane, celkové iónové spektrum (TIS) poskytuje alternatívny prístup k analýze údajov [11]. TIS je totožný s priemerným hmotnostným spektrom pokrývajúcim celý chromatografický rozsah.

Napriek tomu väčšina forenzných laboratórií vykonáva analýzu vzoriek z požiariska prostredníctvom GC-MS. Viacerí autori pracujú na alternatívnych metodikách, ktoré môžu doplniť alebo sa môžu stať alternatívou k tradičným metódam najmä z dôvodu ich výhod.

In this sense, Ferreiro-González *et al.* [7] successfully optimized an electronic nose based on mass spectrometry (HS-MS eNose).

The HS-MS eNose was also applied for analysis of fire debris samples without any pre-treatment [8] and the results were validated by reference method [9].

The alternative HS-MS eNose proved to be useful in the analysis of fire debris with the same discrimination power as GC-MS.

3 Chemometrics

Chemometrics (multivariate statistical analysis) can be used to interpret results of forensic analyses, especially those involving pattern recognition. By using multivariate statistical methods, the statistical significance of the differences in these patterns can be established. When using multivariate statistical techniques, replicate sample measurements should be made to allow for experimental uncertainty and determine the significance of between-sample differences [12,13].

Fortunately, the use of multivariate statistical analysis is a growing practice in forensic chemistry. Forensic scientists often have to identify patterns and interpret differences in data. Chemometrics makes this task more accurate, objective, and manageable. It is especially useful when the scientist is presented with large quantities of data. For example, comparing hundreds of spectra, chromatograms, etc. by inspection was never a valid scientific technique, but was widely used (and sometimes still is) until the adoption of multivariate statistical techniques by increasing numbers of forensic chemists. Multivariate statistics has been used on many types of forensic trace evidence, including accelerants, inks, fibers, smokeless powder, glass, and paint [13,14].

In general, a chemometric approach to a large data set follows a number of distinct steps: Acquire data of the highest quality possible.

- Carry out data preprocessing to remove any systematic (and unimportant) differences between samples.
- Utilize unsupervised (pattern recognition) chemometric methods to visualize

V tomto zmysle Ferreiro-González *et al.* [7] úspešne optimalizovali analytickú techniku - elektronický nos s headspace a hmotnostnou spektrometriou (HSMS eNose) za účelom detekcie zvyškov horľavých kvapalín.

HS-MS eNose bola úspešne aplikovaná pri tepelnej desorpcii zvyškov horľavých kvapalín ako alternatíva k použitiu CS₂ rozpúšťadla [8]. Táto metóda bola úspešne aplikovaná aj na analýzu simulovaných vzoriek z požiariska bez akejkoľvek predprípravy [9].

3 Chemometria

Chemometria (viacrozmerné štatistické metódy) môže byť použitá na interpretáciu výsledkov z forenzných analýz, najmä tých, ktoré zahŕňajú rozpoznávanie špecifických profilov / vzorov. Použitím chemometrie je možné stanoviť štatistickú významnosť premenných a ich rozdielov v študovaných modeloch. Pri použití chemometrie by sa mali vykonať opakované merania vzoriek s cieľom umožniť experimentálne odchýlky a určiť ich význam medzi jednotlivými vzorkami [12,13].

Našťastie použitie chemometrických metód narastá v praxi forenznnej chémie. Forenzní vedci musia často identifikovať špecifické profily a interpretovať ich rozdiely v rámci získaných údajov. Chemometria robí túto úlohu presnejšou a objektívnejšou. Tieto metódy sú obzvlášť užitočné, ak je vedec vystavený veľkému množstvu údajov, napríklad pri analyzovaní chromatografických údajov, spektier, a pod. Vizualne porovnávanie chromatogramov nebolo nikdy platnou vedeckou metodikou, ale bolo široko používané (a niekde stále je) až do implementácie viacrozmerných štatistických metód.

Viacnásobná štatistika sa používa na objasnenie mnohých typov súdnych dôkazov vrátane akcelerátorov, atramentov, vlákien, farbív, bezdymového prášku, skla a farby [13,14].

Všeobecne platí, že chemometrický prístup k veľkej množine údajov pozostáva z viacerých krokov:

- Získanie údajov najvyššej možnej kvality.
- Predbežné spracovanie údajov na odstránenie všetkých systematických (a nedôležitých) rozdielov medzi vzorkami.

- The underlying structure of the data (e.g., the number of classes).
- Utilize supervised (pattern recognition) chemometric techniques to predict the class membership of unknown samples [13,15].

In order to make data interpretation less subjective and avoid masking the presence of ILRs, in recent years chemometric tools and multivariate statistical analysis have been successfully applied to fire debris analysis [16]. In this sense, Hierarchical Cluster Analysis (HCA), Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), as well as Soft Independent Modelling of Class Analogies (SIMCA) have been applied to unburned and to burned ignitable liquid samples [17,18] and Artificial neural networks (ANN) has been applied to gasoline classification due to near-infrared data [19]. Tan *et al.* [17] were successful in classifying 51 liquids into classification system provided by ASTM. Baerncopf *et al.* [20] investigated light and heavy conditions of burn on carpet with six different ignitable liquids. By applying PCA and Pearson product moment correlation (PPMC) coefficients the authors were able to associated with neat ILs.

4 Cluster Analysis

Cluster analysis is considered an unsupervised method in that the algorithm does not rely on any inputs from the user, so no prior knowledge of groupings is known. The purpose of cluster analysis is to determine whether individual samples fall into groupings and what those groupings might be [21].

Hierarchical clustering methods produce a set of nested clusters organized in a hierarchy tree (Fig. 1). The cluster hierarchy is typically visualized using dendrograms. Such approaches are applied either to provide multiresolution data organization or to alleviate computational challenges when clustering big data sets. In general, two approaches are applied to build nested clusters:

- The divisive approach starts with the entire data as one cluster, and then iteratively split clusters until a stopping criterion (e.g., k clusters or tight enough clusters) is satisfied.
- The agglomerative approach starts with small tight clusters, or even with single-

- Aplikácia štatistických metód (bez prvého určenia kategórií) na vizualizáciu základnej štruktúry údajov (napr. počet tried).
- Aplikácia metód na rozpoznávanie vzorov a na predpovedanie členstva v triedach neznámych vzoriek [13,15].

Aby sa interpretácia údajov stala menej subjektívnou a aby sa zabránilo maskovaniu prítomnosti zvyškov horľavých kvapalín, v posledných rokoch došlo k úspešnej aplikácii chemometrických nástrojov a viacrozmerných štatistických analýz v rámci interpretácie dát zo vzoriek z požiariska [16]. V tomto zmysle boli na spálené a nespálené vzorky horľavých kvapalín aplikované analýzy: Hierarchická zhluková analýza (HCA), Analýza hlavných komponentov (PCA), Lineárna diskriminačná analýza (LDA) a SIMCA (Soft Independent Modelling Class Analogie) [17,18]. Umelé neuronové siete (ANN) sa použili na klasifikáciu benzínu prostredníctvom dát pochádzajúcich z infračervených údajov [19]. Tan *et al.* [17] boli úspešní pri klasifikácii 51 horľavých kvapalín do klasifikačných tried určených ASTM. Baerncopf *et al.* [20] skúmali spálený koberec so šiestimi rôznymi horľavými kvapalinami. Použitím PCA a Pearsonových koeficientov korelácie momentu produktu (PPMC) sa autorom podarilo prepojiť spálené zvyšky horľavých kvapalín s ich čistými podobami.

4 Zhluková analýza

Zhluková analýza sa považuje za metódu „bez dozoru“ pretože algoritmus nezávisí od žiadnych vstupných informácií, takže nie sú známe žiadne predchádzajúce znalosti o klastroch (skupín). Cieľom zhlukovej analýzy je určiť, či jednotlivé vzorky spadajú do klastrov a aké sú tieto zoskupenia [21].

Metódy zhlukovej analýzy vytvárajú súbor klastrov usporiadaných v hierarchickom strome (obrázok 1). Hierarchia klastra sa zvyčajne vizualizuje pomocou dendrogramov. Takýto prístup sa uplatňuje pri organizovaní dát s viacerými riešeniami alebo na pri zhrmažďovaní veľkých dátových súborov. Vo všeobecnosti sa na budovanie klastrových zoskupení používajú dva prístupy:

- Rozdeľovací prístup začína celým údajom ako jeden klaster a potom sa iteratívne prerozdeľujú klastre, až kým nie je splnené

point clusters, and then iteratively merge close clusters until only a single one remains [22].

kritérium zastavenia (napr. k počet klastrov alebo tesné usporiadanie klastrov).

- Aglomeratívny prístup začína malými tesnými klastrami, alebo dokonca s jednobodovými klastrami, a potom sa iteračne zhľukujú blízke klastre, kým zostane len jeden [22].

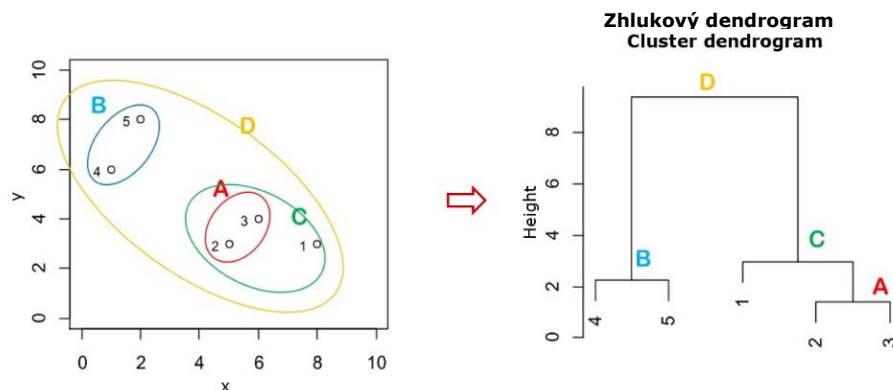


Fig. 1 Hierarchical clustering organized in a hierarchy tree

Obr. 1 Hierarchické zoskupenie zorganizované do hierarchického stromu

5 Principal component analysis

Principal component analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set [23].

In general terms, PCA uses a vector space transform to reduce the dimensionality of large data sets (Fig. 2). Using mathematical projection, the original data set, which may have involved many variables, can often be interpreted in just a few variables (the principal components) by their linear combinations. It is therefore often the case that an examination of the reduced dimension data set will allow the user to spot trends, patterns and outliers in the data, far more easily than would have been possible without performing the principal component analysis [24,25].

The first principal component is the direction along which the samples show the largest variation. The second principal component is the direction uncorrelated to the first component along which the samples show the second largest variation. Each component

5 Analýza hlavných komponentov

Analýza hlavných komponentov (PCA) je matematický algoritmus, ktorý znižuje rozmernosť údajov pri zachovaní väčšiny variácií v súbore údajov [23].

Vo všeobecnosti sa PCA využíva na transformáciu vektorového priestoru a zmenšenie rozmerov veľkých dátových súborov (Obr. 2). Pomocou matematickej projekcie môže byť pôvodný set údajov obsahujúcich veľa premenných, interpretovaná len niekoľkými premennými (hlavnými zložkami) a ich lineárnymi kombináciami. Z toho dôvodu preskúmanie súboru údajov v zmenšenom rozmere umožní používateľovi zistiť trendy, vzory a odchýlky v údajoch oveľa ľahšie, než by bolo možné bez vykonania analýzy hlavných zložiek [24,25].

Prvou hlavnou zložkou je smer pozdĺž ktorého vzorky vykazujú najväčšiu odchýlku. Druhou hlavnou zložkou je smer nekorelovaný k prvému komponentu, pozdĺž ktorého vzorky vykazujú druhú najväčšiu variáciu. Každá zložka môže byť interpretovaná ako smer,

can then be interpreted as the direction, uncorrelated to previous components, which maximizes the variance of the samples when projected onto the component [26].

ktorý nie je v korešpondencii s predchádzajúcimi komponentmi, čo maximalizuje odchýlky vzoriek pri projektovaní na komponent [26].

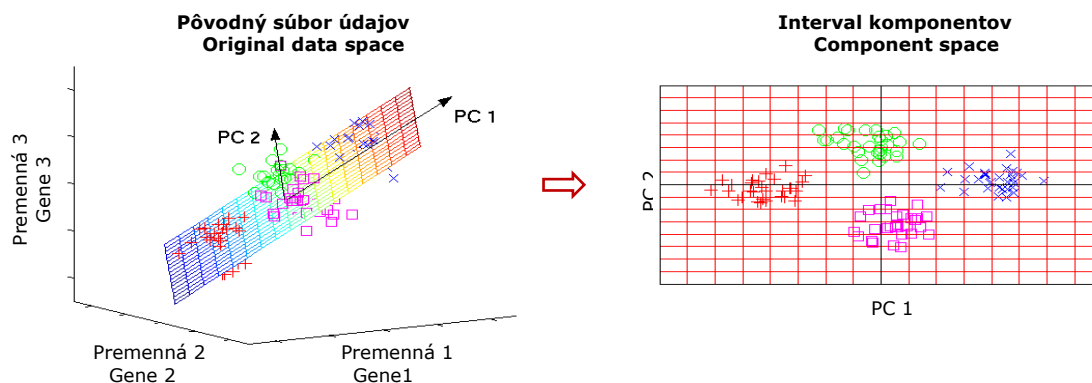


Fig. 2 Dimension reduction from original data set to principal components
Obr. 2 Eliminácia dimenzií z pôvodného súboru dát do hlavných komponentov

6 Discriminant analysis

Discriminant analysis (DA) as a whole is concerned with the relationship between a categorical variable and a set of interrelated variables. In discriminant analysis, the existence of the groups is known a priori. An entity of interest is assumed to belong to one (and only one) of the groups.

Discriminant analysis is widely used also in the field of pattern recognition [27][28]. There are several purposes for discriminant analysis (Fig. 3):

- To classify cases into groups using a discriminant prediction equation.
- To test theory by observing whether cases are classified as predicted.
- To investigate differences between or among groups.
- To determine the most parsimonious way to distinguish among groups.
- To assess the relative importance of the independent variables in classifying the dependent variable.
- To discard variables which are little related to group distinctions [29].

6 Diskriminačná analýza

Diskriminačná analýza (DA) sa ako celok zaoberá vzťahom medzi kategorickou premennou a súborom vzájomne súvisiacich premenných. V diskriminačnej analýze je existencia skupín známa. Predpokladá sa, že skúmaná jednotka patrí k jednej (a iba jednej) skupine.

Diskriminačná analýza je široko používaná aj v oblasti rozpoznávania vzorov / profilov [27][28]. Existuje niekoľko účelov diskriminačnej analýzy (Obr. 3):

- Zatriedenie dát do skupín pomocou diskriminačnej predikčnej rovnice.
- Testovanie teórie tak, že analyzujete, či sú dáta klasifikované tak ako boli predpovedané.
- Preskúmanie rozdielov medzi skupinami alebo v rámci nich.
- Určenie najrelevantnejšieho spôsobu určenia rozdielov medzi skupinami.
- Zhodnotenie relatívnej dôležitosti nezávislých premenných pri klasifikácii závislej premennej.
- Odstránenie premenných, ktoré majú malý vzťah k skupinovým rozdielom [29].

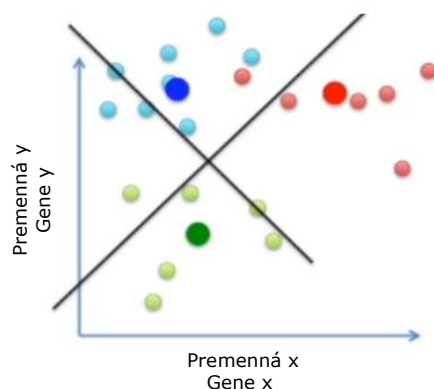


Fig. 3 Axis that maximize the distance among the means for the three categories
Obr. 3 Osi ktoré maximalizujú vzdialenosti medzi stredmi troch skupín

Discriminant analysis has basic two steps: (1) an F test (Wilks' lambda) is used to test if the discriminant model as a whole is significant, and (2) if the F test shows significance, then the individual independent variables are assessed to see which differ significantly in mean by group and these are used to classify the dependent variable.

7 Conclusion

The paper deals with a brief description of multivariate statistical analysis and specify chemometric approach to a data set. Application of chemometrics to fire debris seems to be a promising tool in data interpretation and determination of ignitable liquid residues, as it is more objective and accurate.

Acknowledgments

This work was funded by the projects KEGA 012/TU Z-4/2016 (50 %) and KEGA 030UMB-4/2017 (50 %)

References / Literatúra

1. TURVE, Brent E.; Crowde; Stan. *Forensic Investigations: An Introduction*. Academic Press, 2017. ISBN 978-0-12-800680-1.
2. MARTÍN-ALBERCA, Carlos, ORTEGA-OJEDA, Fernando Ernesto and GARCÍA-RUIZ, Carmen. Analytical tools for the analysis of fire debris. A review: 2008-2015. *Analytica Chimica Acta* [online]. 2016. Vol. 928, p. 1–19. DOI 10.1016/j.aca.2016.04.056.
3. STAUFFER, E., et al. *Fire Debris Analysis*. Boston: Academic Press, 2008. ISBN 978-0126639711.
4. BORUSIEWICZ, Rafał, GRZEGORZ, Zadora and ZIĘBA-PALUS, Janina. Application of

Diskriminačná analýza má dva základné kroky: 1. test F (Wilksova lambda) sa používa na testovanie, či je diskriminačný model ako celok významný a 2. ak test F vykazuje významnosť, potom sa hodnotia jednotlivé nezávislé premenné aby sa zistilo, ktoré sa výrazne líšia v priemere podľa skupín a tieto sa používajú na klasifikáciu závislej premennej.

7 Záver

Príspevok sa zaoberá stručným opisom viac-rozmerných štatistických metód a špecifikuje chemometrický prístup k analýze dát. Aplikácia chemometrie na vzorky z požiariska sa zdá byť sľubným nástrojom pri interpretácii údajov a stanovení zvyškov horľavých kvapalín, z dôvodu vyššej objektivnosti a presnosti.

Pod'akovanie

Táto práca bola financovaná z projektov KEGA 012/TU Z-4/2016 (50 %) a KEGA 030UMB-4/2017 (50 %).

- Head-Space Analysis with Passive Adsorption for Forensic Purposes in the Automated Thermal Desorption-Gas Chromatography-Mass Spectrometry System. *Chromatographia*. 2004. Vol. 60, no. S1, p. 133–142. DOI 10.1365/s10337-004-0299-4.
5. WADDELL, Erin E., WILLIAMS, Mary R. and SIGMAN, Michael E. 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. Vol. 59, no. 4, p. 927–935. DOI 10.1111/1556-4029.12417.
 6. NATIONAL ACADEMY OF SCIENCE. *Strengthening Forensic Science in the United States: A Path Forward*. 2009. ISBN 9780309131353
 7. FERREIRO-GONZÁLEZ M.; AYUSO, J.; ÁLVAREZ, J. A.; PALMA, M.; BARROSO, C. G. New headspace-mass spectrometry method for the discrimination of commercial gasoline samples with different research octane numbers. *Energy & Fuels*. 2014. Vol. 28, no. 10, p. 6249–6254. DOI 10.1021/ef5013775.
 8. FERREIRO-GONZÁLEZ, Marta, BARBERO, Gerardo F., PALMA, Miguel, AYUSO, Jesús, ÁLVAREZ, José A. and BARROSO, Carmelo G. Determination of ignitable liquids in fire debris: Direct analysis by electronic nose. *Sensors (Switzerland)*. 2016. Vol. 16, no. 5. DOI 10.3390/s16050695.
 9. FERREIRO-GONZÁLEZ, Marta, BARBERO, Gerardo F., AYUSO, Jesús, ÁLVAREZ, José A., PALMA, Miguel and BARROSO, Carmelo G. Validation of an HS-MS method for direct determination and classification of ignitable liquids. *Microchemical Journal* [online]. 2017. Vol. 132, p. 358–364. DOI 10.1016/j.microc.2017.02.022.
 10. WILLIAMS, Mary R., SIGMAN, Michael E., LEWIS, Jennifer and PITAN, Kelly McHugh. Combined target factor analysis and Bayesian soft-classification of interference-contaminated samples: Forensic Fire Debris Analysis. *Forensic Science International* [online]. 2012. Vol. 222, no. 1–3, p. 373–386. DOI 10.1016/j.forsciint.2012.07.021.
 11. SIGMAN, M.E., WILLIAMS, M.R., CASTELBUONO, J.A., COLCA, J.G., CLARK, C.D. Ignitable liquid classification and identification using the summed-ion mass spectrum. *Instrum. Sci. Technol.* 2008. DOI 36:375–393.
 12. LAVINE, Barry K and WORKMAN, Jerome. Chemometrics. *Analytical chemistry*. 2012. Vol. 80, no. 12, p. 4519–4531. DOI 10.1021/ac303193j.
 13. MENDLEIN, A., SZKUDLAREK, C. and GOODPASTER, J.V. Chemometrics. *Encyclopedia of Forensic Sciences*. 2013. P. 646–651. DOI 10.1016/B978-0-12-382165-2.00259-2.
 14. BRERETON, Richard G. *Applied Chemometrics for Scientists*. John Wiley & Sons, Ltd., 2007. ISBN 978-0-470-01686-2.
 15. OTTO, M. Chemometrics. *Chemometrics*. 2007. P. 1–11. DOI 10.1002/9783527699377.
 16. MARTÍN-ALBERCA, Carlos, ORTEGA-OJEDA, Fernando Ernesto and GARCÍA-RUIZ, Carmen. Analytical tools for the analysis of fire debris. A review: 2008-2015. *Analytica Chimica Acta*. 2016. Vol. 928, p. 1–19. DOI 10.1016/j.aca.2016.04.056.
 17. TAN, Beijing, HARDY, James K. and SNAVELY, Ralph E. Accelerant classification by gas chromatography/mass spectrometry and multivariate pattern recognition. *Analytica Chimica Acta*. 2000. Vol. 422, no. 1, p. 37–46. DOI 10.1016/S0003-2670(00)01050-3.
 18. SANDERCOCK, P. M.L. and DU PASQUIER, E. Chemical fingerprinting of unevaporated automotive gasoline samples. *Forensic Science International*. 2003. Vol. 134, no. 1, p. 1–10. DOI 10.1016/S0379-0738(03)00081-1.
 19. DOBLE, Philip, SANDERCOCK, Mark, DU PASQUIER, Eric, PETOCZ, Peter, ROUX, Claude and DAWSON, Michael. Classification of premium and regular gasoline by gas chromatography/mass spectrometry, principal component analysis and artificial neural networks. *Forensic Science International*. 2003. Vol. 132, no. 1, p. 26–39. DOI 10.1016/S0379-0738(03)00002-1.
 20. BAERNCOFF, Jamie M., MCGUFFIN, Victoria L. and SMITH, Ruth W. Association of Ignitable Liquid Residues to Neat Ignitable Liquids in the Presence of Matrix Interferences Using Chemometric Procedures. *Journal of Forensic Sciences*. 2011. Vol. 56, no. 1, p. 70–81. DOI 10.1111/j.1556-4029.2010.01563.x.

21. MENDLEIN, A., SZKUDLAREK, C. and GOODPASTER, J.V. Chemometrics. *Encyclopedia of Forensic Sciences* [online]. 2013. P. 646–651. DOI 10.1016/B978-0-12-382165-2.00259-2.
22. SHALIZI, Cosma. Distances between Clustering , Hierarchical Clustering. *Data Mining* [online]. 2009. No. September, p. 36–350.
23. STANGL, Dalene K. Principal Component Analysis. *Journal of the American Statistical Association* [online]. 2002. Vol. 103, no. 482, p. 881–882. DOI 10.1007/b98835.
24. RINGNÉR, Markus and RINGNER, M. What is principal component analysis? *Nat Biotechnol* [online]. 2008. Vol. 26, no. 3, p. 303–304. DOI 10.1038/nbt0308-303.
25. ABDI, Hervé and WILLIAMS, Lynne J. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2010. Vol. 2, no. 4, p. 433–459. DOI 10.1002/wics.101.
26. HATCHER, Larry. Principal Component Analysis. *A Step-By-Step Approach to Using the SAS System for Factor Analysis and Structural Equation Modeling* [online]. 1994. P. 1–56. DOI 10.1002/ibd.21544.
27. HUBERTY CARL J.; OLEJNIK S. 2005 *Applied MANOVA and Discriminant Analysis*. John Wiley & Sons. 2005. 488 p. ISBN: 978-0-471-46815-8
28. MCLACHLAN, Geoffrey J. *Discriminant analysis and statistical pattern recognition*. 2004. Wiley-Interscience. ISBN 9780471691150.
29. GARSON., G. DAVID *Discriminant Function Analysis*. Statistical Associates Publishing, 2012. 142 p. ISBN 9781626380097.