



## Combined cluster and discriminant analysis: An efficient chemometric approach in diesel fuel characterization



Márton Novák<sup>a</sup>, Dóra Palya<sup>a</sup>, Zsolt Bodai<sup>a</sup>, Zoltán Nyiri<sup>a</sup>, Norbert Magyar<sup>b</sup>,  
József Kovács<sup>c</sup>, Zsuzsanna Eke<sup>a,d,\*</sup>

<sup>a</sup> Eötvös Loránd University, Joint Research and Training Laboratory on Separation Techniques (EKOL), 1/A, Pázmány Péter sétány, Budapest 1117, Hungary

<sup>b</sup> Budapest Business School, University of Applied Sciences, Department of Methodology, 9-11, Alkotmány utca, Budapest 1054, Hungary

<sup>c</sup> Eötvös Loránd University, Department of Physical and Applied Geology, 1/C, Pázmány Péter sétány, Budapest 1117, Hungary

<sup>d</sup> Wessling International Research and Educational Center, 56, Főti út, Budapest 1047, Hungary

### ARTICLE INFO

#### Article history:

Received 26 July 2016

Received in revised form 13 November 2016

Accepted 16 November 2016

Available online 23 November 2016

#### Keywords:

Combined cluster and discriminant analysis

Chemometrics

Diesel fuel

Compound-specific isotope analysis

Environmental forensics

### ABSTRACT

Combined cluster and discriminant analysis (CCDA) as a chemometric tool in compound specific isotope analysis of diesel fuels was studied. The stable carbon isotope ratios ( $\delta^{13}\text{C}$ ) of *n*-alkanes in diesel fuel can be used to characterize or differentiate diesels originating from different sources. We investigated 25 diesel fuel samples representing 20 different brands. The samples were collected from 25 different service stations in 11 European countries over a 2 year period. The *n*-alkane fraction of diesel fuels was separated using solid-state urea clathrate formation combined with silica gel fractionation. The stable carbon isotope ratios of C10–C24 *n*-alkanes were measured with gas chromatography–isotope ratio mass spectrometry (GC–IRMS) using perdeuterated *n*-alkanes as internal standards. Beside the 25 samples one additional diesel fuel was prepared and measured three times to get totally homogenous samples in order to test the performance of our analytical and statistical routine.

Stable isotope ratio data were evaluated with hierarchical cluster analysis (HCA), principal component analysis (PCA) and CCDA. CCDA combines two multivariate data analysis methods hierarchical cluster analysis with linear discriminant analysis (LDA). The main idea behind CCDA is to compare the goodness of preconceived (based on the sample origins) and random groupings. In CCDA all the samples were compared pairwise.

The results for the parallel sample preparations showed that the analytical procedure does not have any significant effect on the  $\delta^{13}\text{C}$  values of *n*-alkanes. The three parallels proved to be totally homogenous with CCDA.

HCA and PCA can be useful tools when the examining of the relationship among several samples is in question. However, these two techniques cannot be always decisive on the origin of similar samples. The initial hypothesis that all diesel fuel samples are considered chemically unique was verified by CCDA. The main advantage of CCDA is that it gives an objective index number about the level of similarity among the investigated samples. Thus the application of CCDA supplemented by the traditionally used multivariate methods greatly improves the efficiency of statistical analysis in the CSIA of diesel fuel samples.

© 2016 Elsevier Ireland Ltd. All rights reserved.

## 1. Introduction

Middle distillate fuel oils such as diesel fuel are frequently spilled in the environment. Those cases when the source of the spills is questionable or totally unknown are providing serious challenges in environmental forensic investigations. In order to determine the liability associated with the cleanup and remediation chemical fingerprinting methods are applied. In the case of source correlation studies of unknown fuel contaminations the evaluation of similarities or dissimilarities among samples or among a sample and a possible source is the problem to be solved.

\* Corresponding author at: Eötvös Loránd University, Joint Research and Training Laboratory on Separation Techniques (EKOL), 1/A, Pázmány Péter sétány, Budapest 1117, Hungary.

E-mail addresses: [marton.novak@ekol.chem.elte.hu](mailto:marton.novak@ekol.chem.elte.hu) (M. Novák), [dora.palya@ekol.chem.elte.hu](mailto:dora.palya@ekol.chem.elte.hu) (D. Palya), [zsolt.bodai@ekol.chem.elte.hu](mailto:zsolt.bodai@ekol.chem.elte.hu) (Z. Bodai), [zoltan.nyiri@ekol.chem.elte.hu](mailto:zoltan.nyiri@ekol.chem.elte.hu) (Z. Nyiri), [Magyar.Norbert@uni-bge.hu](mailto:Magyar.Norbert@uni-bge.hu) (N. Magyar), [kevesolt@geology.elte.hu](mailto:kevesolt@geology.elte.hu) (J. Kovács), [eke.zsuzsanna@wirec.eu](mailto:eke.zsuzsanna@wirec.eu) (Z. Eke).

Therefore there are two types of question in spill studies. The first type arises if there is a presumed contamination source and a contaminated sample and the question is if they are the same or not. The second type is when there are several contaminated samples and there is not necessarily a hypothetic source so the relationship and the grouping of the samples are in question.

The correlation between the samples is usually buried in a complex data set, thus the application of multivariate statistical methods is indispensable in almost all cases. The literature of chemometric techniques in oil spill fingerprinting is quite extensive thus it is not discussed here only the literature relevant in diesel fuel fingerprinting is considered. A comprehensive review published by Christensen and Tomasi [1] is available focusing on the practical aspects of chemometric techniques in fingerprinting studies.

Fingerprinting studies of diesel fuels are primarily focused on GC–MS analysis of the composition of compound groups affording a high discriminatory potential among fuels. These compound groups are *n*-alkanes, isoprenoids, alkylcyclohexanes, alkylbenzenes, alkylated polycyclic aromatic hydrocarbons (alkyl-PAHs) and low molecular weight biomarkers such as sesquiterpanes, adamantanes and diadamantanes [2–6]. Only a few studies are available focusing on the appropriate chemometric techniques for diesel fuel fingerprinting based on GC–MS data. Marshall et al. [7] compared five neat diesel samples based on total ion chromatograms (TICs) and four alkane and four aromatic extracted ion chromatograms (EICs) using principal component analysis (PCA) and Pearson product moment correlation (PPMC). Hupp et al. [8] analyzed 25 diesel samples using PCA and PPMC. The comparison was based on the TIC and *m/z* 57, 91 and *m/z* 141 EICs of the samples. Gaines et al. [9] investigated 14 diesel fuel samples using PCA. For every sample peaks were integrated in 62 EICs representing 22 compound classes and 313 peak ratios were generated. Nine peak ratios including that of alkylbenzenes, alkylphenantrenes and sesquiterpanes were found by PCA which are necessary to differentiate the diesel samples from one another. Ebrahimi and Hibbert [10] compared a diesel fuel sample from an oil spill and seven suspected source samples which were subject to controlled weathering using multi-way parallel factor analysis (PARAFAC2). The comparison was based on a set of 10 *m/z* ratios of two- and three-ring PAHs.

Compound-specific isotope analysis (CSIA) has also proven its utility in middle distillate fingerprinting. The prevalent absence of the routinely used diagnostically relevant components (e.g. high molecular weight biomarkers) and the possible impact of weathering on the above mentioned and usually measured compound classes ensure the significance of stable isotope ratio determinations. Despite the special value of stable isotope ratio data only a few studies are available for the characterization of diesel fuel by CSIA [11–15]. In addition fingerprinting based on stable isotope ratio data have been lacking a statistical routine which is appropriate for objective decision making when an unknown contaminant and a possible source are compared as well as for the examining of relationship among several contaminated samples.

Out of these few studies only the one published by Boyd et al. [13] focuses on the importance of multivariate statistical tests in CSIA of oil spill studies. They performed sample-wise PCA on the stable carbon isotope ratios of *n*-alkanes in diesel fuel contaminations then performed a multivariate analysis of variance (MANOVA) upon the principal component variables. Hierarchical cluster analysis (HCA) was also used in the statistical processing routine as an additional method to look for relationships among the investigated samples. Boyd et al. addressed the problem that multivariate statistical tests require high numbers of parallel measurements and in the case of CSIA the high cost and the time

demand of a measurement makes it fairly impractical to perform the sufficient amount of replicate analysis. Another study where diesel fuels were differentiated based on stable isotope ratio data of *n*-alkanes was published by Muhammad et al. [15]. They applied the widely used statistical methods PCA and HCA for the differentiation of 45 diesel fuel samples from 9 different suppliers in the South Island of New Zealand.

With regard to the previously cited two publications it is important to note that the application of chemometric methods is valuable only if the validity of the data thus the validity of the analytical method has been clearly assessed. In the case of CSIA peak purity is a key factor in the validity of data. The absence of attention on data quality may lead to erroneous data interpretation and may mislead the scientist if the performance of the chemometric approaches is in question. However, the previously cited studies did not address the importance of peak purity in CSIA. The only one study where selective sample preparation – i.e. solid-state urea clathrate formation – was used to separate *n*-alkanes from the complex diesel fuel matrix was published by Harvey et al. [14]. In this study the differentiation of four diesel fuel samples from different locations in the USA was reported. Stable isotope ratio data were assessed by PCA. The samples were clearly distinguished from one another. However, the samples were totally different of origin and only four diesel fuels were analyzed. Hence the literature of CSIA in oil spill forensics is still lacking a study where a high number of samples are analyzed including totally homogenous, similar and different samples and special attention is paid on peak purity.

The aim of this research was to apply an efficient chemometric method, the combined cluster and discriminant analysis (CCDA) in the characterization of diesel fuels based on stable isotope ratio data of *n*-alkanes. This multivariate data analysis method was developed and published by Kovács et al. [16] and successfully applied in environmental studies in Tanos et al. [17] and Kovács et al. [18]. We also aimed to work out a statistical routine in the CSIA of diesel fuels which is suitable for objective decision making when an unknown contaminant and a possible source are compared as well as for the examining of relationship among several samples. To this end, we applied PCA and HCA as additional tools to utilize all the advantages and capabilities of the different multivariate statistical methods simultaneously. We decided to collect and analyze a high number of diesel fuel samples including similar and totally different samples in order to test the performance of our multivariate statistical routine. Special attention was paid on data quality of stable isotope ratio determinations. CSIA of *n*-alkanes was carried out using a chemical separation method based on silica gel fractionation combined with solid state urea clathrate formation [19]. We also aimed to solve the problem of the amount of data demanded by multivariate statistical evaluation and the amount of parallel measurements obtainable by CSIA within reasonable limits.

## 2. Materials and methods

### 2.1. Chemicals and reagents

*n*-Hexane was obtained from Sigma-Aldrich (Steinheim, Germany) and was of at least 99% purity. Methanol was purchased from LGC Promochem GmbH (Wesel, Germany) and was of LiChrosolv® quality. Urea was obtained from Molar Chemicals Ltd. (Budapest, Hungary). Silica gel cartridges (Isolute SI 500 mg/3 mL with 50 µm particle size and 60 Å nominal porosity) were obtained from Biotage (Uppsala, Sweden). Nitrogen (purity 4.5) was from Messer Group GmbH (Bad Soden, Germany). The deuterated *n*-alkanes: *n*-dodecane-*d*<sub>26</sub>, *n*-hexadecane-*d*<sub>34</sub>, *n*-tetracosane-*d*<sub>50</sub>

Download English Version:

<https://daneshyari.com/en/article/6462525>

Download Persian Version:

<https://daneshyari.com/article/6462525>

[Daneshyari.com](https://daneshyari.com)