



Chemical characterization of the acid alteration of diesel fuel: Non-targeted analysis by two-dimensional gas chromatography coupled with time-of-flight mass spectrometry with tile-based Fisher ratio and combinatorial threshold determination



Brendon A. Parsons^a, David K. Pinkerton^a, Bob W. Wright^b, Robert E. Synovec^{a,*}

^a Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98198, USA

^b Pacific Northwest National Laboratory, Battelle Boulevard, P.O. Box 999, Richland, WA 99352, USA

ARTICLE INFO

Article history:

Received 16 December 2015

Received in revised form 19 February 2016

Accepted 22 February 2016

Available online 27 February 2016

Keywords:

Tile-based F-ratio

GC × GC–TOFMS

Null distribution analysis

Diesel fuel

Acid alteration

False discovery rate

ABSTRACT

The illicit chemical alteration of petroleum fuels is of keen interest, particularly to regulatory agencies that set fuel specifications, or taxes/credits based on those specifications. One type of alteration is the reaction of diesel fuel with concentrated sulfuric acid. Such reactions are known to subtly alter the chemical composition of the fuel, particularly the aromatic species native to the fuel. Comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC × GC–TOFMS) is well suited for the analysis of diesel fuel, but may provide the analyst with an overwhelming amount of data, particularly in sample-class comparison experiments comprised of many samples. Tile-based Fisher-ratio (F-ratio) analysis reduces the abundance of data in a GC × GC–TOFMS experiment to only the peaks which significantly distinguish the unaltered and acid altered sample classes. Three samples of diesel fuel from differently branded filling stations were each altered to discover chemical features, i.e., analyte peaks, which were consistently changed by the acid reaction. Using different fuels prioritizes the discovery of features likely to be robust to the variation present between fuel samples and may consequently be useful in determining whether an unknown sample has been acid altered. The subsequent analysis confirmed that aromatic species are removed by the acid alteration, with the degree of removal consistent with predicted reactivity toward electrophilic aromatic sulfonation. Additionally, we observed that alkenes and alkynes were also removed from the fuel, and that sulfur dioxide or compounds that degrade to sulfur dioxide are generated by the acid alteration. In addition to applying the previously reported tile-based F-ratio method, this report also expands null distribution analysis to algorithmically determine an F-ratio threshold to confidently select only the features which are sufficiently class-distinguishing. When applied to the acid alteration of diesel fuel, the suggested per-hit F-ratio threshold was 12.4, which is predicted to maintain the false discovery rate (FDR) below 0.1%. Using this F-ratio threshold, 107 of the 3362 preliminary hits were deemed significantly changing due to the acid alteration, with the number of false positives estimated to be about 3. Validation of the F-ratio analysis was performed using an additional three fuels.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Petrochemical fuels are refined by a variety of methods that allow for inherently variable feedstocks to be better controlled in finished products to meet a range of quality metrics and other specifications, such as engine performance and longevity, storage

stability, and emissions. Such processes and specifications ensure that the fuels are acceptable for distribution and end use by consumers. However, between the refinery and end user, intermediate parties may further alter the chemical composition of finished fuels, for the purpose of altering the fuel color [1]. Hence, there is substantial interest in the forensic detection of such alteration by law enforcement and regulatory agencies. An alteration that has been observed in the field is the reaction of petroleum fuels with concentrated sulfuric acid (H₂SO₄). Sulfuric acid is a prevalent and inexpensive reagent used in a wide variety of industrial,

* Corresponding author. Fax: +1 206 685 8665.

E-mail address: synovec@chem.washington.edu (R.E. Synovec).

commercial, and household applications. Based on the chemical composition of petroleum fuels, one would expect reaction with sulfuric acid to cause a variety of chemical changes in fuel due to reactions between the sulfuric acid and native compounds in the fuel [2–4]. Gasoline fuel exhibits substantial chemical changes due to the presence of *tert*-butylating compounds, such as methyl *tert*-butyl ether (MTBE), while the changes to diesel fuel are more subtle, limited to the aromatic moiety of the fuel [2,4]. For the purposes of forensic classification of acid altered diesel fuel, it is necessary to better understand the chemical compositional changes that occur and which may differentiate unaltered from altered fuels.

Reactions between sulfuric acid and aromatic compounds are well represented in the literature, particularly in the fields of surfactant and dye chemistry, in which electrophilic species in the sulfuric acid react with the aromatics to add sulfonic acid, altering the aqueous solubility of the aromatics [5]. Literature on sulfonation reactions using sulfuric acid indicate that various aromatic species have an intrinsically different reactivity with the electrophilic species in the concentrated acid, and thus will be chemically modified to different extents as the reaction proceeds [6–8]. An interesting aspect of aromatic sulfonation by sulfuric acid addition is that the reaction produces water, leading to a quenching of the reaction as concentrations of the electrophilic species decrease. Most studies on the reaction rates and extent of reactions for aromatic sulfonation using sulfuric acid involve simple reagent systems (i.e., neat reagent or reagent dissolved in a suitable solvent) and an excess of sulfuric acid (typically 3–4 molar equivalents). In the forensic study of acid altered diesel fuel, these ideal conditions are absent: first, the speciation of aromatics in diesel is highly complex, often ranging from benzene to pyrene with various degrees of alkylation and unsaturate substitution, as well as lesser amounts of aromatic heterocycles. Additionally, the sulfuric acid is probably not added stoichiometrically, and factors such as the water content of the diesel fuel may also impact the rates and extent of reaction. Further, industrial syntheses typically include temperature control and removal of water from the sulfuric acid to assist in driving the aromatic sulfonation to completion; such controls may not be present in illicit alterations.

Due to the substantial differences expected between the literature reaction conditions for synthesis of sulfonated aromatics versus the less-ideal conditions for illicit alteration of diesel fuel, a non-targeted “discovery-based” analysis would be beneficial for characterizing the chemical changes that result from the alteration. In non-targeted analysis, the chemical features (i.e., analyte peaks) of interest are not presumed to be known prior to the experiment. Non-targeted approaches may be performed using supervised or unsupervised methods, where supervision refers to the prior classification of samples as they relate to the experimental design [9–11]. Overall, non-targeted methods aim to discover the key chemical features, i.e., the analyte peaks indicative of chemical marker compounds that robustly distinguish samples from each other. For the purpose of analyzing the chemical composition of diesel fuel, comprehensive two-dimensional (2D) gas chromatography coupled with time-of-flight mass spectrometry (GC \times GC–TOFMS) provides the necessary analytical resolution, sensitivity, and sample throughput [12–19]. Recently, a comprehensive and informative review was reported, dealing with the application of GC \times GC in forensics [20]. Indeed, GC \times GC is increasingly being utilized in forensic studies, but it is acknowledged that the complexity of the data is a major factor in slowing the technology’s implementation. To improve the utility of GC \times GC for all users, and to increase its utilization in forensic applications, it is necessary to develop feature selection software to reduce the complex output to only the most significant chemical features.

We previously reported the development of tile-based Fisher ratio (F-ratio) analysis for the comprehensive discovery-based

analysis of GC \times GC–TOFMS data [21,22]. This methodology, herein referred to as tile-based F-ratio analysis, is a supervised non-targeted method that discovers features in a sample-class comparison (e.g., unaltered fuel versus acid altered fuel) that are class-distinguishing, i.e., peaks that are statistically different between sample classes. The tiling portion of the software utilizes a novel binning scheme which improves the discovery of class-distinguishing chemical features (“hits”) while minimizing the discovery of features which are not significantly different between sample classes (“non-hits”). Tile-based F-ratio analysis coupled with GC \times GC–TOFMS is highly effective for the experimental determination of chemical changes in the highly complex matrix of diesel fuel.

In the process of feature selection using tile-based F-ratio analysis, the chromatograms for the sample-class comparison experiment are comprehensively analyzed, with only the most essential application of signal thresholds, enabling the discovery of class-distinguishing analytes, even at the trace concentration level. The F-ratio, which is the calculated value for the one-way analysis of variance (ANOVA) *F*-test [23] is used to rank chromatographic features in a hit list according to their class-distinguishing abilities, with a large F-ratio indicating that a given feature is highly class-distinguishing, and a small F-ratio indicating that a given feature is not class-distinguishing. In order to avoid false positives (type I errors) and to utilize analyst time most efficiently, only features with a sufficiently high F-ratio should be pursued in subsequent analyte deconvolution, identification and quantification steps; this may be accomplished by setting an F-ratio threshold for the hit list. However, determination of such a threshold is often a subjective matter that requires manual analyst intervention to decide the point at which the selected features become unreliable, and often leads to excessive, time-consuming manual data analysis steps [24,25].

As an alternative to manual threshold determination, statistics-based approaches should be considered and applied. One approach is to utilize the critical value for the *F*-test (F_{crit}) to determine a threshold for significance [26,27]. Another statistics-based approach is using a z-scores method [28]. An alternative approach for F-ratio threshold determination is to analyze the within-class variation for each sample class within the experiment, and utilize the maxima of these values as the cutoff, under the basis that any significant between-class variation will exceed that of the within-class variation [29]. We recently reported a new approach, null distribution analysis, which is a permutation-based test [30] to establish the distribution of spurious F-ratio values for a given GC \times GC–TOFMS experimental design analyzed using tile-based F-ratio analysis [22], which used a frequentist strategy to infer a threshold to differentiate hits from non-hits. Herein, we describe improvements to null distribution analysis to determine a statistically-appropriate F-ratio threshold aimed at minimizing false discoveries while simultaneously maximizing true discoveries. The improved null distribution analysis methodology is then applied to find class-distinguishing analytes, i.e., chemical marker compounds which discriminate acid altered from unaltered diesel fuel.

2. Theory

For robust determination of class-distinguishing analytes in non-targeted GC \times GC–TOFMS F-ratio experiments, it is necessary to choose an appropriate threshold to cut off the analysis of the resulting hit list, in order to save analyst time and to avoid false discoveries. Most traditional statistical approaches would suggest simply using a critical value, in this case F_{crit} , which is determined for the degrees of freedom in the experiment and significance level

Download English Version:

<https://daneshyari.com/en/article/7610119>

Download Persian Version:

<https://daneshyari.com/article/7610119>

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