Classification of Biodiesel and Fuel Blends using Gas Chromatography - Differential Mobility Spectrometry with Cluster Analysis and Isolation of C18:3 me by Dual Ion Filtering

by

Dedeepta Pasupuleti, Gary A. Eiceman, and Karisa M. Pierce*

Department of Chemistry and Biochemistry
New Mexico State University
Las Cruces, NM 88003

*Department of Chemistry and Biochemistry
Seattle Pacific University
3307 Third Avenue West, Suite 205
Seattle, WA 98119
piercek@spu.edu

March 17, 2016

ABSTRACT

Fatty acid alkyl esters (FAAEs) were determined at 10 to 100 mg/L in biodiesel and blends with petrodiesel without sample pre-treatment using gas chromatography with a tandem differential mobility detector. Selectivity was provided through chromatographic separations and from atmospheric pressure chemical ionization reactions in the detector with mobility characterization of gas ions. Limits of detection were ~0.5 ng with average repeatability in peak area of 2.98% RSD, in retention time of ≤ 1.3 % RSD, and in compensation voltage of $\leq 9.2\%$ RSD. Biodiesel blends could be classified using principal component analysis (PCA) and hierarchical cluster analysis (HCA). Unsupervised cluster analysis captured 52.72% of variance in a single PC while supervised analysis captured 71.64% of variance using Fisher ratio feature selection. Test set predictions showed successful clustering according to source or feedstock when regressed onto the training set model. Detection of the regulated substance methyl linolenate (C18:3 me) was achieved in 6 to 10 s with a 1 m long capillary column using dual ion filtering in the tandem differential mobility detector.

Key Words: Gas chromatography – dynamic mobility spectrometry; fatty acid alkyl ester; biodiesel/petrodiesel blend; feedstock classification; principal component analysis

INTRODUCTION

Biodiesel has a prominent role in the development of alternative fuels as worldwide energy demand rises. The annual production of biodiesel in 2009 was 4.16 billion gallons worldwide, which was distributed in Europe (49.8%), North and South America (32.8%), and Asia Pacific (4.4%) [1, 2]. By 2020, global production is anticipated to be ~12 billion gallons [1]. Since biodiesel feedstock includes diverse vegetable and animal sources within an expanding marketplace, regulations have been developed for the variability in quality of biodiesel and blends with petrodiesel. The composition of biodiesel and biodiesel/petrodiesel blends is regulated in Europe under DIN EN 14214 and in the United States under ASTM D6751 where a principal specification is the content of fatty acid alkyl ester (FAAE) with alkyl groups commonly methyl (me) or ethyl (et). The total amount of FAAE is related to fuel performance and the eight isomers of methyl linolenate (or C18:3 me), with an 18-carbon chain and 3 double bonds, and a specific determinant of fuel quality. The level of C18:3 me in biodiesel is regulated and minimized to improve fuel performance.

Several methods have been reported for quantitative or semi-quantitative determination of FAAEs in biodiesel and for chemometric classification of biodiesel feedstock. Those include optical spectroscopic methods [3-5], NMR [6], various chromatographic methods, commonly with mass spectrometry (MS) detectors [7, 8], MS alone [9], and MS with specialized ionization techniques [10]. A significant challenge for analysis of biodiesel/petrodiesel blends for FAAEs using gas chromatography (GC), or GC-MS, is the relatively abundant and chemically complex matrix of hydrocarbons. Selectivity has been introduced through sample pre-treatment to isolate FAAEs and through advanced separation methods based on GC×GC [11] and GC-MS [12] such as selected ion monitoring (SIM) with GC-MS measurements [13] and suppression of

matrix interferences using chemical ionization at ambient pressure (APCI) with MS methods [14].

An alternative to complex instrumentation is ion mobility spectrometry (IMS) where selectivity in response begins with APCI reactions and is supplemented through ion characterization by mobility in air at ambient pressure. These features favor hyphenated-combinations with gas chromatographs and a version of IMS, known as differential mobility spectrometry (DMS), was incorporated into a portable field-rugged GC-DMS analyzer with column temperature programming [15]. The benefits of selectivity in response and a second dimension of analytical information with DMS as a GC detector was shown recently with high speed GC-DMS analysis of biodiesel and biodiesel-petrodiesel blends [16]. Selectivity of the detector allowed methyl linolenate to be determined in less than 10 seconds (s) with a tandem version of DMS, where an additional layer of selectivity was introduced through dual ion filtration. In principle, rapid and convenient measurements of biodiesel or biodiesel/petrodiesel blends by GC-DMS methods could be made on-site anywhere in the production and distribution cycle. Although these exploratory findings were promising, data handling was manually intensive and classification through chemometric methods was unproven. Critically, quantitative performance was undetermined and classification of authentic samples using computational methods, which rely upon repeatability, was not demonstrated.

Chemometric tools to classify or analyze biodiesels for chemical information have included principal component analysis (PCA) [17, 18], PCA-projections to latent structures [14], partial least squares (PLS) [19, 20], PLS combined with PCA, hierarchical cluster analysis (HCA), K-nearest neighbor (KNN) [21], and support vector machines (SVM) with PLS/ multivariate adaptive regression splines (MARS) [22]. Other applications of chemometric methods include quantitative determinations of blends with biodiesel, petrodiesel, and a renewable hydrocarbon by NIR spectrometry [4, 23], the percent composition in biodiesel/petrodiesel by GC×GC-TOFMS [11], evaluation of blend quality by proton NMR [24], classification of blends by FT-ICR

MS [22], and the determination of methyl esters in biodiesel blends using FTIR-ATR and FTNIR [25]. Chemometric tools have been applied successfully with complex data sets from GC-DMS analyses although matrix interferences were not considered significant complications [15].

Thus, chemometrics may be effective with GC-DMS measurements depending on the robustness of ion formation in the complex composition of petrodiesel.

The objectives in the present work were to explore the quantitative response of GC-DMS toward FAAEs in standard solutions and authentic samples and to explore the effectiveness of chemometric methods to classify biodiesel feedstocks. Dual ion filtering, developed for tandem DMS measurements previously to isolate C18:3 content, was demonstrated with synthetic and authentic biodiesel blends.

EXPERIMENTAL

Instrumentation

A tandem differential mobility spectrometer was coupled with a 5890 series gas chromatograph (Hewlett-Packard Corp, Avondale, PA) creating a GC-tandem DMS instrument. The gas chromatograph was equipped with a split-splitless injector and a HP-5 capillary column, 30 m long $\times$ 0.25 mm ID (Agilent Technologies Inc., Santa Clara, CA). The column effluent was introduced into the differential mobility spectrometer through a 0.5 m non-polar aluminum clad capillary column inside a 25 cm long $\times$ 6 mm OD (1/8th inch) stainless steel tube at 220°C. Carrier gas was nitrogen filtered through molecular sieves and also a Model No: OT3-2 oxygen/moisture trap (Agilent Technologies, Palo Alto, CA). The injector split ratio was 25:1 and carrier gas velocity was $\sim$ 30 cm/s.

The tandem DMS drift tube was from ChemRing Detection Systems (Charlotte, NC, USA) and was based on two DMS stages sequentially placed with a separation distance of 2 mm between the first (DMS1) and second (DMS2) stages. Electronics to control the tandem DMS detector

were adapted from two control boards of JUNO®, a hand-held DMS analyzer, and specialized software was provided by ChemRing Detection Systems to operate the tandem DMS instrument. The ion source was a 2 mCi ⁶³Ni foil located inside a SAF 2507® Super Duplex Stainless compression fitting union (Swagelok, El Paso, TX, USA). The carrier gas was supplemented with 0.5 L/min of air purified through 13x molecular sieve to a moisture level of 1 ± 0.5 ppm determined with a Moisture Image Series 2 (Panametrics, Inc. Waltham, MA, USA). The drift tube was held in a PEEK frame providing gas tight connections through high temperature resistant fluorocarbon O rings (part no: 20377, Restek Corp.) for flow into and from the drift tube. The PEEK frame was encased in a heated aluminum shell at 85°C unless noted otherwise. Separation voltages (V) of 600, 700, 800, 900, and 1000 corresponded to separation fields (Townsend, Td) of 67.5, 78.7, 90 Td, 101.2, and 112.5, respectively. The tandem DMS electronics provide a separation voltage as high as 1495 V or a field of 168.7 Td. Barometric pressure in Las Cruces, NM was nominally 660 mmHg. Methylene chloride was introduced as a vapor modifier into the gas flow for the DMS drift tube. Headspace over a liquid reservoir, in a glass vial at 30°C, was passed at 0.5 L/min into 2.2 L/min of purified air yielding a vapor concentration, determined gravimetrically, of 1% (wt/vol). Gas flows were adjusted using mass flow controllers (Sierra Instruments, Inc., Monterey, CA). An HP model 5890 gas chromatograph with HP 5971 Mass Selective Detector was used for GC/MS characterization of samples using scanning mode. The GC was equipped with the same column used for GC-DMS studies and a split-split less inlet. Carrier gas was helium with split ratio 25:1.

Reagents and Samples

Sixteen FAAE standards were obtained from Sigma Aldrich (St. Louis, MO). Hexane (HPLC grade, Fisher Scientific) was used to prepare dilutions of the esters. Dilutions of the esters ranged from 10 to 500 ppm. Methylene chloride (HPLC grade, Fisher Scientific) was used as a

vapor modifier in the DMS. Methylene chloride was used as a modifier in the gas atmosphere to enhance ion peak separation for ions with masses of 150 Da or larger. Such ions in a purified gas atmosphere exhibit relatively flat alpha functions and appear near 0 Td in compensation field (CF) showing little or no resolution. In the presence of vapor modifiers such as isopropanol, water, and methylene chloride at 3% w/v, differences in mobility coefficients of an ion are increased and ion peaks are displaced to greater values of |CF| [26, 27]. From a local commercial vendor in Seattle, WA, a 100% biodiesel sample was obtained. Samples of two different 100% jatropha biodiesels (designated jatropha I and jatropha II) and two different 100% soy biodiesels (designated soy A and soy B) were obtained from an energy technology company (Richmond, CA). A sample of unaltered petrodiesel was obtained from Restek Corporation (Bellefonte, PA). Blends of biodiesel with petrodiesel were also obtained from a local fuel vendor in Las Cruces, NM, and these blends purportedly contained biodiesel at the following volume percentages: 0%, <5%, and 5 to 20%. New Mexico State University students from the Environmental Engineering department produced and provided what was purportedly 100% biodiesel from algae feedstock.

Procedures

GC-tandem DMS analysis with the tandem DMS detector in single stage mode:

The first stage, DMS1, was set to all-ion pass mode and the separation field (SF) on the second stage, DMS2, was set to 111.4 Td. The compensation field (CF) on DMS2 was continuously swept from -1.13 to 0.9 Td at a step size of 0.03 Td with 15 ms per step. One microliter of a mixture of 16 FAAEs at 250 ppm each in hexane was analyzed and results were processed in Microcal ORIGIN 9.0, MATLAB R2009b by Math-Works, and PLS Toolbox Version 6.2.1 by Eigenvector Research, Inc (Manson, WA). Chromatography parameters were: Injection port, 215°C; initial temperature, 150°C; final temperature, 262°C; and temperature program rate, 3°C/min.

GC-tandem DMS analysis with Selected Ion Peak Monitoring:

In these measurements, values for SF and CF for both stages, DMS 1 and DMS 2, were fixed to pass ions which met the criteria of pairs of SF: CF in both DMS stages. A window of ~ 0.2 Td was allowed on the CF value for DMS 2. This provided dual filtering of ions continuously with effluent from the gas chromatograph.

Gas Chromatography-Mass Spectrometry:

In addition to GC-DMS analyses, we decided to also analyze our samples by GC-MS to confirm compound identities using relative retention times and mass spectra, and to evaluate chromatographic parameters. Chromatography parameters were as follows: injection port, 225°C; initial temperature, 150°C; final temperature, 265°C; temperature program rate, 3°C/min; transfer line, 200°C; and splitless injection time, 0.5 min. One microliter amounts of sample were analyzed.

High Speed analysis by GC-DMS and GC-tandem DMS:

The chromatographic column used above was replaced with 2.5 m long HP 5 column joined to the 0.5 m aluminum nonpolar transfer column. Column temperature was isothermal at 270°C and the drift tube was 120°C. Solutions of C16:0 me, C18:0 me, C18:1-cis-9 me, C18:2 me and C18:3 me were individually analyzed by GC-DMS with compensation field scans from -0.89 to 1.10 Td at a step size of 0.03 Td. Mixtures of these five esters at various concentrations were analyzed under the same conditions. Biodiesels from different feedstock were diluted with solvent to various concentrations and were analyzed under the same conditions. For the tandem DMS measurements, specific values for separation and compensation fields on each stage, (i.e., SF1:CF1 and SF2:CF2) were chosen to selectively pass the C18:3 me ester through dual ion filtering.

Quantitative Studies using GC-DMS:

Four standard mixtures containing four FAAEs in each were prepared and further diluted with hexane to six different concentrations from 10 to 250 ppm. Each mixture was measured by GC-DMS in triplicate. One mixture contained C8:0 me, C14:0 me, C18:2 cis-9,12 me, and C18:2 cis-9,12 et; a second mixture had C10:0 me, C18:1 cis-9 me, C18:1 cis-9 et, and C20:0 me; a third mixture had C15:0 me, C18:3 cis 9,12,15 me, C18:3 cis 9,12,15 et, and C18:0 et; and a fourth mixture had C16:0 me, C10:0 et, C18:1 cis 6 me, and C18:0 me. The GC-tandem DMS parameters were: DMS1 set to off; DMS 2 set to SF = 101.2 Td and CF sweeping from -1.1 Td to + 0.9 Td at 0.03 Td step size.

Procedures for Unsupervised Clustering and Supervised Classification of Fuels Using GC-DMS:

The GC-DMS chromatograms of 41 samples for which the feedstocks were known were acquired and combined into a *training set*. The GC-DMS chromatograms of seven samples for which the feedstocks were known were acquired and combined into a *test set*. The GC-DMS chromatograms of nine commercial samples were also acquired and the feedstocks of these samples were truly unknown, so these were combined into an unknown *prediction set*.

The GC-DMS chromatograms were exported as Excel files from the instrument software. Then the pixel level GC-DMS chromatograms were imported into MATLAB. In MATLAB, each GC-DMS chromatogram had a smoothed blank sample chromatogram subtracted from it to yield a baseline-corrected chromatogram. Then each baseline-corrected chromatogram was unfolded along the compensation field dimension and aligned using the peak-match algorithm with window size equal to 6 points [28]. The alignment target was the first jatropha sample chromatogram from the training set; it was chosen because it contained more peaks than many of the other chromatograms in the training set. This same target was noted and used to align all later test set and prediction set chromatograms, as well. Each baseline-corrected, unfolded,

and aligned GC-DMS chromatogram was then scaled. The scaling was achieved by dividing each signal that was greater than a noise threshold (0.015 units) by the maximum signal in the chromatogram, causing each scaled GC-DMS chromatogram to have a final maximum signal that was equal to 1.0. The purpose of this scaling was to remove variations in blend percentage while aiming to maintain and preserve any FAAE profile variations that were capable of differentiating feedstocks. The preprocessed training set chromatograms were combined into a single 2D matrix, and if the analysis was unsupervised, then at this point the matrix was simply mean centered and submitted to HCA (agglomerative HCA with Ward's method dendrogram) or PCA to build the HCA and PCA models. If the analysis was supervised, then right after being combined into the single 2D matrix, this was the point at which the supervised Fisher ratio algorithm was applied to select only the features that discriminated the feedstocks in the training set. Then only those selected features from the training set/test set/prediction set were subsequently mean centered using the training set mean, and submitted to HCA and PCA [29]. The preprocessed test set chromatograms were regressed onto the training set models to allow the analyst to predict the feedstock of each test set sample. The nine truly unknown prediction set samples were also regressed onto the models to predict feedstock.

RESULTS AND DISCUSSION

Gas Chromatography of Fatty Acid Alkyl Esters with MS and DMS Detectors

Gas Chromatography-Mass Spectrometry of Biodiesel Blends

Chromatographic performance was determined using GC-MS as a referee method and retention times for FAAEs are given in Table 1, establishing elution order and relative retention for later GC-DMS studies with the same stationary phase. Results from GC-MS analysis of unaltered petrodiesel sample from Restek (0% biodiesel) are shown as a total ion chromatogram (TIC) in

Figure 1A and the principle constituents, from analysis of mass spectra, were alkanes, alkyl substituted benzenes, and a large undetermined number of minor or trace constituents. Major constituents at retention times between 3 min and 15 min were n-alkanes with carbon number from C10 to C22 (Figure 1A). Of the more than 60 major or significant constituents, fatty acid alkyl esters (FAAEs) were not detected in the petrodiesel sample based on intensities of the extracted ion chromatogram for ion 74.1 Da, the base peak for methyl esters (Figure 1B). Analysis of a mixture of petrodiesel, obtained from a Las Cruces vendor, and 16 FAAEs (yielding a 20% (v/v) biodiesel/petrodiesel blend) produced the TIC in Figure 1C and extracted ion chromatogram for 74.1 Da in Figure 1D. Comparison of Figure 1A to Figure 1C shows that simple inspection alone of the TIC under these conditions is insufficient to readily disclose the presence of FAAEs in this blend of petrodiesel and biodiesel, even at the 20% (w/w) of 16 FAAEs (12.5 ng per FAAE), though this was possible with an extracted ion chromatograms (Figure 1D). Findings showed suitable separation of FAAEs with carbon numbers C8 to C16 with this 30 m HP-5 stationary phase column yet separations with carbon numbers C18 or greater were incomplete and chromatographic conditions could not be found for complete or even partial separation of some FAAEs. These findings demonstrate that chromatographic separations with a routine stationary phase might benefit from another dimension of separation, particularly for the C18 isomers. A differential mobility detector alone or in tandem DMS mode of operation may provide this dimension with portability for on-site measurements.

Gas Chromatography-Differential Mobility Spectrometry of Biodiesel Blends

Ion mobility methods with chemical ionization sources at ambient pressure in air or nitrogen (APCI) are distinguished by selectivity from reactions between substances and gas phase hydrated protons (as shown in Equation 1). Alkanes, substituted alkanes and some aromatic hydrocarbons, without dipoles to retain a proton through electrostatic attraction, will not

yield product ions or exhibit response as shown in Figure 2A from GC-DMS analysis of the petrodiesel sample. In comparison to the total ion chromatogram from GC-MS analysis with electron impact ionization (Figure 1A), the simplicity of the chromatogram in Figure 2A suggests that the use of GC-DMS measurements in chemometric analyses will not be complicated by hydrocarbons in the diesel fraction of biodiesel blends. A hidden benefit of combining a gas chromatograph with a APCI based method such as DMS is the improved reliability and simplicity of response when substances are introduced into the ion source individually or simple mixtures and ionization chemistry is uncomplicated by competitive reactions.

In the GC-DMS chromatogram of a standard solution of FAAEs, separations resembled those from GC-MS for peaks 1 to 6 (Figure 2B); however, improved chromatographic efficiency was seen for peaks 7 to 16 although this column was the same length and stationary phase as used in the GC-MS measurements. Even so, two unresolved groups of peaks were seen for compounds 7 to 10 and 12 to 14. The result for a petrodiesel blend with authentic biodiesel sample (20% soy biodiesel) showed comparable chromatographic results with selectivity from APCI reactions (Figure 2C). Five FAAEs (6, 7, 8, 10 & 11) are observed without any signal contributions from petrodiesel constituents. This removal of component signal from petrodiesel hydrocarbons from the component signals of FAAEs demonstrates significant simplicity in chemometric analysis of biodiesel/petrodiesel blends.

Determination of FAAEs in Synthetic Biodiesel Blends through GC-DMS Analysis

The other benefit of a differential mobility detector is the additional dimension of information available from the mobility characterization of ions. This additional step transforms DMS from a simple one dimensional ionization detector into a multidimensional detector that is capable of low limits of detection as well as providing molecular information about substances as ions.

Such content is seen Figure 3A in a topographic plot of ion intensity (z axis), retention time (minutes), and compensation field (Td) from the GC-DMS separation of a mixture of 16 FAAEs in methylene chloride as solvent. In Figure 3B for a 20% (w/w) mixture of FAAEs with Las Cruces petrodiesel, the peaks with prominent response were FAAEs and additional substances were detected as peaks with low abundances throughout the elution time. In Figure 3C, results from an authentic sample of biodiesel/petrodiesel blend (purportedly <5%) obtained from a Las Cruces vendor exhibited ion peaks for five FAAEs and a larger abundance of other substances throughout the elution profile. These substances were at trace levels and yet were distinctive of only this sample. Findings in Figures 1A, 2A, and 2C showed that, for the most part, GC-DMS was insensitive to alkanes, alkenes, and aromatic hydrocarbons (though some response to polycyclic aromatic hydrocarbons in IMS is known) [30]. Though these peaks were not detectable in GC-MS analysis of the exact same sample, relative elution times match molar masses larger than the bulk of constituents in unaltered petrodiesel and consequently, could be less volatile substances that accumulate over time as the fuel reacts with air or microbes. Substances responsible for such response could include oxygen- or nitrogen- containing substances, principally aromatic substances. It is possible that petrodiesel obtained from the local Las Cruces vendor would be contaminated with such substances during normal transfer and storage events. In the following discussions, we will refer to these components of the local fuel vendor's petrodiesel as "unknown constituents" whenever the sample of interest contained the Las Cruces vendor's petrodiesel.

Quantitative Performance

Reliable quantitative response from GC-DMS analyses is necessary for chemometric processing of measurements and APCI sources are known to exhibit limited linear ranges and source saturation [31]. Results from measurements of standards are shown in Figure 4 and Table 2. Response curves show onset of response near 0.04 ng and a linear response at and

above 40 ng. During this range and with the low amounts of analyte and large excess of reactant ions, the reaction in Equation 1 is pseudo first order and linear in response, as shown in Equation 2.

$$\text{Rate} = k' [M]$$

2

where $k' = k [H^+(H_2O)]$. At higher amounts of analyte, the rate of production of reactant ions is insufficient for the reaction with M proceeding near the rate of collision and the rate equation becomes second order; the response curve is non-linear. In studies here, the source was not saturated and results were confined to this linear range.

The precision of quantitative response is shown in Table 2 with peak areas, standard deviation of peak area, and percent relative standard deviation (% RSD) of peak area are listed for each FAAE. The average standard deviation (and range) of peak area was 0.032 (0.002 to 0.088) with an average %RSD (and range) of 2.98% (0.071% to 25.2%). The precision of commercial handheld analyzers range from 3 to 10 % RSD of peak area, so the sources of quantitative variation in provided by the GC-DMS instrument was regarded as sufficient. Precision of retention time as RSD was measured largely at 1% and lower and only a few are reported here: methyl caprate with 1.3% RSD, ethyl oleate with 0.54% RSD, and methyl arachidate with 0.31% RSD. Similarly, the precision of compensation voltage as % RSD was ~9% and lower with examples of methyl caprate, ethyl oleate, and methyl arachidate of 9.2, 3.3, and 1.7%, respectively. The compensation field precision for FAAE peaks in petrodiesel matrix suggests the CF values for the FAAEs were unaffected by the presence of the unknown constituents in the local vendor's petrodiesel. These levels of precision for retention time and CF suggest that the GC-DMS data are quantitatively suitable for chemometric modeling. Rather, a key question from these findings is: can a tandem DMS detector provide sufficient information from

chromatographic data sets to be algorithmically assessed and used to predict the feedstock from which each biodiesel was originally derived?

Unsupervised Cluster Analysis and Classification of Blends of Biodiesel and Petrodiesel Using GC-DMS

Blends of petrodiesel with variable feedstocks of biodiesel (two sources of soy, two sources of jatropha, and a commercial biodiesel) were submitted to unsupervised and supervised classification analysis to determine whether GC-DMS could provide the information necessary to predict the biodiesel feedstock of unknown biodiesel/petrodiesel blends. Results from analysis of several samples are shown in Figure 5 where differences in composition (regarding peaks detected and relative abundances) were seen in extracted ion intensity at CF = 0.59 Td. Unsupervised PCA of GC-DMS responses of unaltered petrodiesel and biodiesel/petrodiesel blends are shown in Figure 6A (filled circles for training set members and filled triangles for test set members). Test set chromatograms were regressed onto the training set model in order to predict the feedstocks of the test set members. The percent variances captured were PC 1 = 52.73%, PC 2 = 13.71%, and PC 3 = 7.37%. Every test set sample was nearest to samples in the training set that originated from the same feedstock. Thus, we claim that PCA objectively clustered the test set samples according to the correct feedstock. However, for discussion purposes, we subjectively and manually drew four boundaries around the training set scores that belonged to the same/similar feedstock. Thus, according to the training set scores in Figure 6A, four main clusters were detected in the unsupervised PCA analysis of the training set. One of those clusters consisted solely of biodiesel/petrodiesel blend samples from both jatropha feedstocks which were clustered together and this cluster was separate from all other feedstocks. A second cluster seen in the training set consisted solely of both soy biodiesel feedstocks which were clearly clustered together and the cluster was separate from all other feedstocks. Likewise, a third cluster consisted solely of "Seattle commercial" biodiesel

feedstock. The training set also consisted of two other kinds of fuels which, according to both our GC-DMS and GC-MS chromatograms, contained no observable FAAEs (purportedly “algae biodiesel” produced and provided by university students) and the unaltered petrodiesel standard from Restek Corp. The scores of these two kinds of fuels (Restek petrodiesel and purportedly algae biodiesel blended with petrodiesel) were clustered closely together in a fourth cluster which was distant from the other three clusters. To avoid having to rely on predictions made using manually subjective hand-drawn boundaries, an HCA dendrogram was built for the same training set and it is shown in Figure 6B. The clustering in this dendrogram is consistent with the PCA plot, especially if a vertical line is drawn at approximately 10 units of variance weighted distance (vwd) between cluster centers. That is the greatest distance at which the four clusters are correctly delineated, and correct clusters are predicted from approximately 8 to 10 vwd units. In this HCA plot, the nearest neighbors of all test set members are the correct group of training set members with matching feedstock. Thus it is arguable that both PCA and HCA correctly clustered the test set members according to feedstock. The training set’s success at correctly predicting the test set’s feedstocks fostered some confidence in the PCA and HCA models. It is noteworthy that the unsupervised clustering of all 41 training set samples mostly matched feedstock sources even though the biodiesel percent volume composition of these biodiesel/petrodiesel blends varied from 5% to 20% (v/v) within each feedstock. In other words, the entire analytical process described herein, so far, was able to correctly predict feedstock variations despite variations in the ratio of biodiesel to petrodiesel within each class of feedstock. This success was partly due to the scaling technique described in the Procedures section; when the scaling was not applied, then the clustering did not correspond with feedstock groups. The degree-of-class-separation between each feedstock cluster is reported in Table 4.

The HCA and PCA models were then used to predict the truly unknown samples that had been set aside in a “prediction set”. These prediction set samples were composed of

biodiesel/petrodiesel blends that were obtained from a local vendor in Las Cruces, NM, and we did not know the original feedstock of these biodiesels. The vendor claimed that these samples were biodiesel blended with petrodiesel to levels of (purportedly) 0%, <5%, and 5 to 20% (v/v) biodiesel in the blend. As shown in Figure 7, PCA and HCA clustered these truly unknown prediction set samples into a new fifth cluster, so we conclude these samples might be from a feedstock that is different than our soy, jatropha, or Seattle commercial feedstock. We were surprised the <5% and 5 to 20% samples clustered with the 0% samples in the prediction set. But upon further inspection of both GC-DMS and GC-MS chromatograms, unknown constituents (which were described earlier in the discussion of Las Cruces vendor samples and Figure 3) were present in all of these particular samples which dominated any small FFAE peaks in these purported biodiesel/petrodiesel blends. The PCA and HCA models were able to automatically detect these similarities and place the prediction set members into a separate cluster. The origins of these samples were undocumented and uncontrolled, so these results should not be over-interpreted. What is notable is that our analytical method was able to overcome the variations in blend percent composition and instead successfully focus on feedstock variation. Another way to interpret these prediction set results is they also demonstrate a major limitation of using the PCA and HCA models to predict the feedstock of biodiesel blends, that limitation being these particular models will not be useful for feedstocks that are not included in the training set model. For example, if a biodiesel blend made from animal fat was part of the test set and if it was submitted to the PCA and HCA models, the animal biodiesel's score would probably appear outside the four clusters and the analyst would have to classify it as "other". So PCA and HCA models are able to predict feedstocks only of types that were used to construct the model. In our case, we were only able to acquire five biodiesel samples with trusted feedstock origin information. This is a major limitation of the PCA and HCA methods described herein, but it is conceivable that analysts with access to more

trusted feedstock origins could extend the PCA and HCA methods described herein to model more types of feedstocks.

Supervised Cluster Analysis and Classification of Blends of Biodiesel and Petrodiesel Using GC-DMS

Next, efforts were made to improve clustering by adding a Fisher ratio method supervision step prior to HCA and PCA. A Fisher ratio is a measure of class-to-class variation divided by within-class variation for each data point in a data set of GC-DMS chromatograms which contains data from a variety of sample classes. The Fisher ratio threshold was defined to be 10 times the average standard deviation of noise regions of the GC-DMS chromatograms. We selected and retained only features that had Fisher ratios greater than this threshold in order to remove noise and other within-class variations that could obscure between-class variations during PCA and HCA. The results of submitting the selected features to HCA and PCA are shown in Figure 8. The percent variances captured in PCA were PC 1 = 71.64%, PC 2 = 13.62%, and PC 3 = 5.93%. This supervised PCA model captured a greater percentage of variation in the first three PCs than the unsupervised PCA model. The four clusters are much “tighter” than they were in Figure 6, and the degree-of-class-separation among the clusters, as reported in Table 4, improved with this supervised feature-selection step added to the analytical procedure. Indeed, Figure 8 clearly predicts the distinct class membership of the truly unknown test set samples. It is noteworthy that supervised HCA classification of the training set is clear from approximately 4 to 10 vwd units and HCA objectively and correctly predicted the feedstock for every test set member.

Fisher ratios plot reveals that FAAE profiles provided by GC-DMS can differentiate feedstocks:

In Figure 9, a plot of Fisher ratios as a function of CF and retention time is shown for the 41-member training set. The Fisher ratios plot contains peak-like shapes with retention time and

CF values that match those of FAAEs, corroborating that the entire proposed analytical procedure preserves and detects enough appropriate chemical information to differentiate biodiesel feedstocks based on the unique and reproducible FAAE profiles of the original feedstock materials.

GC-Tandem DMS to isolate methyl linolenate (C18:3 me) from various feedstocks:

In addition to feedstock classification, another goal of this work was fast chromatographic determination of C18:3 me, a key constituent of biodiesel for which the quantitative content is officially regulated. The intention was that dual filtering action of a tandem DMS based on characteristic patterns of dispersion plots would allow the C18:3 me to be isolated and supported against a referee method [18]. Prepared for these studies were 20% (v/v) biodiesel/petrodiesel blends for each feedstock we used in the training set, mixed with unaltered petrodiesel from Restek. The blends were submitted to GC-DMS and GC-tandem DMS. The chemical selectivity provided by GC-DMS and GC-tandem DMS allowed isolation of C18:3 me in less than 10 s of analysis time. Results in Table 3 showed peak area precision of 0.5% RSD to 2.8% RSD.

CONCLUSIONS

Application of GC-DMS to biodiesel/petrodiesel blends demonstrated several benefits of DMS as a detector for FAAEs including: selectivity of ion formation providing simple chromatographic results since alkanes and aromatic hydrocarbons exhibit limited response with chemical ionization reactions at ambient pressure. The GC-DMS provided adequate retention time precision, CF precision, dynamic range, and limits of detection for successful unsupervised and supervised PCA and HCA analysis of FAAEs in petrodiesel matrix. As for the peak area precision provided by GC-DMS specifically for these challengingly massive FAAE components in petrodiesel matrices, our prior publication described how unknown unresolved matrix

components will contribute to the signals of any overlapping FAAEs [18]. A GC-DMS-based method that seeks to quantify the absolute concentration of FAAE components in biodiesel/petrodiesel blends would require further physical resolution, mathematical resolution, and/or multivariate calibration, which we are continuing to work on by introducing detector modifications and partial least squares analysis to the process in future work. To be clear, in this current article, instead of focusing on determining absolute concentration values of FAAEs, herein we focused on determining whether or not the entire process of making biofuel, blending biofuel, and analyzing blends with GC-DMS preserved enough chemical information about the original feedstock. Indeed, we showed that the GC-DMS can provide chromatograms that are a reproducible and unique fingerprint of each blend and these patterns can be algorithmically assessed and used to predict the feedstock from which each biodiesel originally derived.

Acknowledgements:

Financial support from the National Science Foundation, Award No. CHE-1306388 is gratefully acknowledged.

Table 1: Properties and abbreviations of FAAEs that were mixed in-house to synthesize biodiesel. (me = methyl ester and et = ethyl ester).

	Fatty Acid Alkyl Ester (FAAE)	Molar Mass (g/mol)	Retention Time (min)	Density (g/ml)
1	Methyl Caprylate, C8:0 me	158.2	5.5	0.877
2	Methyl Caprate, C10:0 me	186.3	7.8	0.871
3	Ethyl Caprate, C10:0 et	200.3	9.0	0.865
4	Methyl myristate, C14:0 me	242.4	17.3	0.855
5	Methyl penta decanoate, C15:0 me	256.4	20	0.862
6	Methyl palmitate, C16:0 me	270.4	23.8	0.852
7	Methyl Linoleate, C18:2 cis-9,12 me	294.5	29	0.889
8	Methyl Oleate, C18:1 cis-9 me	296.5	29.3	0.874
9	Methyl Petroselinatate, C18:1 cis 6 me	296.5	29.3	0.874
10	Methyl Linolenate, C18:3 cis 9,12,15 me	292.5	29.5	0.895
11	Methyl Stearate, C18:0 me	298.5	30	solid
12	Ethyl Linoleate, C18:2 cis-9,12 et	308.5	31.1	0.876
13	Ethyl Oleate, C18:1 cis-9 et	310.5	31.3	0.87
14	Ethyl Linolenate, C18:3 cis 9,12,15 et	306.5	31.4	0.892
15	Ethyl Stearate, C18:0 et	312.5	32	solid
16	Methyl Arachidate, C20:0 me	326.5	36	solid

Table 2. Average peak area and concentration of 16 standard FAEs. Also included are standard deviations and percent errors for the triplicate measurements at each concentration. Concentrations maintained in obtaining these quantitative plots are 40 ng, 10ng, 4 ng, 2 ng, 0.8 ng and 0.4 ng per microliter injected. Peak areas were determined using plots of ion intensity with respect to retention time at a compensation field of 167 Td.

Mass- 0.4 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	0.369	0.556	1.291	0.260	0.561	0.589	0.662	0.216	0.571	0.266	0.419	0.271
Sd	0.051	0.020	0.019	0.009	0.027	0.031	0.043	0.002	0.053	0.008	0.016	0.049
%RSD	13.87	3.655	1.497	3.494	4.728	5.319	6.468	0.920	9.202	2.893	3.849	17.903

Mass- 0.8 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	0.522	1.016	1.778	0.658	1.070	0.769	0.783	0.369	1.012	0.383	0.697	0.500
Sd	0.023	0.018	0.037	0.020	0.012	0.028	0.018	0.045	0.034	0.005	0.087	0.052
%RSD	4.408	1.747	2.106	3.086	1.080	3.669	2.294	12.133	3.367	1.278	12.441	10.485

Mass- 2 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	0.722	2.191	4.309	0.830	2.578	1.855	2.086	2.588	1.745	0.495	1.193	1.193
Sd	0.027	0.041	0.035	0.030	0.010	0.044	0.054	0.004	0.031	0.030	0.055	0.006
%RSD	3.720	1.894	0.819	3.606	0.396	2.353	2.603	0.163	1.750	6.024	4.621	0.526

Mass- 4 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	1.310	2.814	5.027	1.554	4.818	2.836	3.758	4.952	1.934	1.562	2.243	3.799
Sd	0.054	0.040	0.017	0.027	0.022	0.037	0.010	0.004	0.015	0.003	0.022	0.072
%RSD	4.089	1.415	0.347	1.707	0.457	1.311	0.268	0.071	0.770	0.163	0.969	1.900

Mass- 10 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	2.546	3.807	12.111	3.601	6.536	7.834	5.426	7.184	4.457	2.425	4.071	5.602
Sd	0.050	0.069	0.011	0.038	0.029	0.021	0.033	0.016	0.025	0.032	0.010	0.048
%RSD	1.959	1.821	0.087	1.045	0.450	0.271	0.602	0.227	0.571	1.326	0.248	0.860

Mass- 40 ng

	C8 me	C10:0 me	C10:0 et	C14:0 me	C15:0 me	16:0 me	C18:0 me	C18:0 et	C18:1 et	C18:2 me	C18:2 et	C20 me
Area	5.851	7.039	15.872	6.845	7.700	14.982	13.647	11.387	7.251	5.055	8.614	9.859
Sd	0.050	0.041	0.020	0.033	0.042	0.033	0.014	0.033	0.029	0.052	0.013	0.088
%RSD	0.856	0.588	0.125	0.485	0.544	0.223	0.103	0.294	0.395	1.020	0.148	0.890

Table 3. Peak area precision for C18:3 me (peak 10) in biodiesel/petrodiesel blends provided by tandem -DMS. Algae and jatropha samples showed no or below detection limits responses for C18:3 me (peak 10).

Feedstock	Peak Area in triplicate measurements	Average	Standard deviation	% Relative Standard Deviation
Soy A	0.0431	0.0421	0.0011	2.6
	0.0410			
	0.0421			
Soy B	0.0351	0.0351	0.0002	0.6
	0.0349			
	0.0353			
Seattle commercial	0.0264	0.0273	0.0008	3.0
	0.0279			
	0.0276			

Table 4: Degree-of-class-separation (DCS) obtained from unsupervised (Figure 6A) and supervised (Figure 8A) analysis.

Degree-of-class separation between feedstock clusters	Unsupervised	Supervised
Algae (purportedly)-Seattle commercial	4.28	12.11
Algae (purportedly)-Jatropha	8.43	13.04
Algae (purportedly)-Soy	6.87	10.21
Seattle commercial-Jatropha	1.75	3.21
Seattle commercial-Soy	2.08	3.80
Jatropha-Soy	2.28	3.03

LIST OF FIGURES

1. GC-MS chromatograms plotted as A. total ion current (TIC) from unaltered petrodiesel (Restek Corp); B. selected ion chromatogram (SIC) of m/z 74.1 Da in petrodiesel, C. TIC of 20% (v/v) total standard mixture of 16 FAAEs in petrodiesel, and D. SIC of 74.1 Da of 20% (v/v) blend of 16 FAAEs in petrodiesel.

2. GC-DMS chromatograms plotted as ion intensity at $CF = 0.6$ Td for A. unaltered petrodiesel, B. 10 ng/ μ L each FAAE in methylene chloride, and C. mixture of Soy A biodiesel blended with unaltered petrodiesel. (Peaks are numbered and identified as described in Table 1).

3. GC-DMS chromatograms plotted as ion intensity as a function of retention time and CF for A. total 20%(v/v) of 16 standard FAAEs in methylene chloride solution, B. 20%(v/v) 16 FAAEs blended in Las Cruces vendor petrodiesel, and C. purportedly <5% biodiesel/petrodiesel blend oacquired from Las Cruces vendor. (Peaks are numbered and identified as described in Table 1).

4. GC-DMS response curves for fatty acid alkyl esters. The FAAEs in plot A are: 1. ethyl caprate, C10:0 et, 2. methyl stearate, C18:0 me, 3. methyl caprate, C10:0 me, 4. methyl pentanoate, C15:0 me, 5. methyl palmitate, C16:0 me, 6. methyl caprylate, C8:0 me, 7. methyl myristate, C14:0 me, 8. ethyl stearate, C18:0 et. The FAAEs in plot B are 9. methyl oleate, C18:1 cis-9 me, 10. ethyl oleate, C18:1 cis -9 et, 11. ethyl linoleate, C18:2 et, 12. ethyl linolenate, C18:3 et, 13. methyl arachidate, C20:0 me, 14. methyl linoleate, C18:2 me, 15. methyl petroselinate, C18:1 cis-6 me, 16. methyl linolenate, C18:3 me.

5. GC-DMS chromatograms plotted as extracted ion intensity at CF = 0.59 Td of A. unaltered petrodiesel, B. 20% (v/v) soy A biodiesel/petrodiesel blend, C. 20% (v/v) jatropha biodiesel/petrodiesel blend, and D. 20% (v/v) Seattle commercial biodiesel/petrodiesel blend.

6. Unsupervised modeling of training set and model validation using independent test set. 6A. PCA model showing the 7-member test set (red triangles labeled 42-48) correctly clustering when regressed on to the unsupervised 41-member training set model (filled circles labeled 1-41). 6B. An unsupervised HCA dendrogram modeled using the 41-member training set correctly clusters the 7 members of the test set (labeled 42-48) according to feedstock. In all cases, the training set members were composed of: 1-3 = 20% (v/v) algae biodiesel/petrodiesel blend, 4-6 = 5% (v/v) algae biodiesel/petrodiesel blend, 7-9 = 20% (v/v) Seattle commercial biodiesel/petrodiesel blend, 10-12 = 5% (v/v) Seattle commercial biodiesel/petrodiesel blend, 13 = Restek unaltered petrodiesel, 14-17 = 20% (v/v) jatropha I biodiesel/petrodiesel blend, 18-19 = 5% (v/v) jatropha I biodiesel/petrodiesel blend, 20-22 = 20% (v/v) jatropha II biodiesel/petrodiesel blend, 23-24 = 5% (v/v) jatropha II biodiesel/petrodiesel blend, 25-26 = 20% (v/v) jatropha II biodiesel/petrodiesel blend, 27 = 5% (v/v) jatropha II biodiesel/petrodiesel blend, 28-30 = 20% (v/v) soy A biodiesel/petrodiesel blend, 31-34 = 5% (v/v) soy A biodiesel/petrodiesel blend, 35-37 = 20% (v/v) soy B biodiesel/petrodiesel blend, 38-41 = 5% (v/v) soy B biodiesel/petrodiesel blend. The test set members were composed of: 42 = 5% (v/v) algae biodiesel/petrodiesel blend, 43 = 20% (v/v) Seattle commercial biodiesel/petrodiesel blend, 44 = 5% (v/v) Seattle commercial biodiesel/petrodiesel blend, 45 = Restek unaltered petrodiesel, 46 = 5% (v/v) jatropha II biodiesel/petrodiesel blend, 47 = 20% (v/v) jatropha I biodiesel/petrodiesel blend, 48 = 20% (v/v) soy A biodiesel/petrodiesel blend.

7. Unsupervised model predictions of feedstock for members of prediction set. 7A. PCA model showing the 9-member prediction set (red triangles labeled 42-50) clustered into a separate, unique cluster when regressed on to the unsupervised 41-member training set model (filled circles labeled 1-41). This is the same training set that was described in the caption for Figure 6). 7B. An unsupervised HCA dendrogram modeled using the 41-member training set clustered the 9 members of the prediction set (labeled 42-50) into a separate, unique cluster, suggesting this unknown feedstock is different than the feedstocks that were used in the training set. The training set members (labeled 1-41) are the same as described in the caption for Figure 6. The prediction set members were composed of: 42-44 = <5% (v/v) Las Cruces vendor biodiesel/petrodiesel blend, 45-47 = Las Cruces vendor petrodiesel, 48-50 = 5-20% (v/v) Las Cruces vendor biodiesel/petrodiesel blend.

8. Supervised modeling of training set and model validation using independent test set. 8A. PCA model showing the 7-member test set (red triangles labeled 42-48) correctly clustering when regressed on to the supervised 41-member training set model (filled circles labeled 1-41). 8B. An unsupervised HCA dendrogram modeled using the 41-member training set correctly clusters the 7 members of the test set (labeled 42-48) according to feedstock. These are the same training set and test set that were described in the caption for Figure 6).

9. Supervised model predictions of feedstock for members of prediction set. 9A. PCA model showing the 9-member prediction set (red triangles labeled 42-50) clustered into a separate, unique cluster when regressed on to the supervised 41-member training set model (filled circles labeled 1-41). 9B. A supervised HCA dendrogram modeled using the 41-member training set clustered the 9 members of the prediction set (labeled 42-50) into a separate, unique cluster.

These are the same training set and prediction set that were described in the caption for Figure 7).

10. (A) Plot of Fisher ratios for the 41-member training set wherein four classes were designated: soy A and soy B were considered a single class, jatropha I and jatropha II were considered a single class, Seattle commercial biodiesel was a class, and purported algae and Restek unaltered petrodiesel were combined into a single class. (B) Variables which have Fisher ratios greater than threshold = 24 are plotted in gray-to-black. Variables which have Fisher ratios less than the threshold are plotted in white.

REFERENCES

- [1] Y.C. Sharma, B. Singh, Renewable and Sustainable Energy Reviews 13 (2009) 1646-1650.
- [2] M.A.A. Silva, R.A. Correa, M.G.d.O. Tavares, N.R.A. Filho, Fuel 143 (2015) 16-20.
- [3] L.C. Gontijo, E. Guimarães, H. Mitsutake, F.B.d. Santana, D.Q. Santos, W.B. Neto, Fuel 117 (2014) 1111-1114.
- [4] J.C.L. Alves, R.J. Poppi, Analyst 138 (2013) 6477-6487.
- [5] J. Prakash, A.K. Mishra, Fuel 108 (2013) 351-355.
- [6] M.R. Monteiro, A.R.P. Ambrozin, L.M. Lião, A.G. Ferreira, Fuel 88 (2009) 691-696.
- [7] M.K. Jennerwein, M. Eschner, T. Gröger, T. Wilharm, R. Zimmermann, Energy and Fuels 28 (2014) 5670-5681.
- [8] N.G.S. Mogollon, F.A.d.L. Ribeiro, M.M. Lopez, L.W. Hantao, R.J. Poppi, F. Augusto, Analytica Chimica Acta 796 (2013) 130-136.
- [9] R.G.D. Prates, R. Augusti, I.C.P. Fortes, Energy and Fuels 24 (2010) 3183-3188.

- [10] P.V. Abdelnur, S.A. Saraiva, R.R. Catharino, M. Coelho, N. Schwab, C.M. Garcia, U. Schuchardt, V.d. Souza, M.N. Eberlin, *Journal of the Brazilian Chemical Society* 24 (2013) 946-952.
- [11] K.M. Pierce, S.P. Schale, *Talanta* 83 (2011) 1254-1259.
- [12] K.J. Johnson, S.L. Rose-Pehrsson, R.E. Morris, *Energy and Fuels* 18 (2004) 844-850.
- [13] R.C.M. Faria, M.J.C. Rezende, C.M. Rezende, A.C. Pinto, *Quimica Nova* 30 (2007) 1900-1905.
- [14] I. Eide, K. Zahlisen, *Energy and Fuels* 21 (2007) 3702-3708.
- [15] S. Prasad, K.M. Pierce, H. Schmidt, J.V. Rao, R. Güth, S. Bader, R.E. Synovec, G.B. Smith, G.A. Eiceman, *Analyst* 132 (2007) 1031-1039.
- [16] D. Pasupuleti, K.M. Pierce, G.A. Eiceman, *Journal of Chromatography A* 1421 (2016) 162-170.
- [17] G.W.B. Silva, A.A. Gomes, P.d. Silva, G.B. Costa, D.D.S. Fernandes, M.M. Fontes, G. Veras, *Journal of the American Oil Chemist's Society* 89 (2012) 1165-1171.
- [18] A.M. Hupp, L.J. Marshall, D.I. Campbell, R.W. Smith, V.L. McGuffin, *Analytica Chimica Acta* 606 (2008) 159-171.
- [19] G.F. Ghesti, J.L. de Macedo, V.S. Braga, A.T.C.P. de Souza, V.C.I. Parente, E.S. Figueredo, I.S. Resck, J.A. Dias, S.C.L. Dias, *Journal of the American Oil Chemists' Society* 83 (2006) 597-601.
- [20] Z. Mustafa, R. Milina, P.A. Simeonova, S.L. Tsakovski, V.D. Simeonov, *Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances and Environmental Engineering* 50 (2015) 72-80.
- [21] S.P. Schale, T.M. Le, K.M. Pierce, *Talanta* 94 (2012) 320-327.
- [22] W.F.C. Rocha, B.G. Vaz, G.F. Sarmanho, L.H.C. Leal, R. Nogueira, V.F. Silva, C.N. Borges, *Analytical Letters* 45 (2012) 2398-2411.
- [23] R.E. Morris, M.H. Hammond, J.A. Cramer, K.J. Johnson, B.C. Giordano, K.E. Kramer, S.L. Rose-Pehrsson, *Energy and Fuels* 23 (2009) 1610-1618.
- [24] M.R. Monteiro, A.R.P. Ambrozini, M. da Silva Santos, E.F. Boffo, E.R. Pereira-Filho, L.M. Lião, A.G. Ferreira, *Talanta* 78 (2009) 660-664.
- [25] J.S. Oliveira, R. Montalvão, L. Daher, P.A.Z. Suarez, J.C. Rubim, *Talanta* 69 (2006) 1278-1284.
- [26] N.S. Krylova, E.V. Krylov, J.A. Stone, G.A. Eiceman, *Journal of Physical Chemistry* 107 (2003) 3648-3654.

- [27] G.A. Eiceman, E.V. Krylov, N.S. Krylova, E.G. Nazarov, R.A. Miller, *Analytical Chemistry* 76 (2004) 4937-4944.
- [28] K.J. Johnson, B.W. Wright, K.H. Jarman, R.E. Synovec, *Journal of Chromatography A* 996 (2003) 141-155.
- [29] K.M. Pierce, J.C. Hoggard, J.L. Hope, P.M. Rainey, A.N. Hoofnagle, R.M. Jack, B.W. Wright, R.E. Synovec, *Analytical Chemistry* 78 (2006) 5068-5075.
- [30] S.H. Kim, F.W. Karasek, S. Rokushika, *Analytical Chemistry* 50 (1978) 152-155.
- [31] G.B. Smith, G.A. Eiceman, M.K. Walsh, S.A. Critz, E. Andazola, E. Ortega, F. Cadena, *Field Analytical Chemical and Technology* 4 (1997) 213-226.

Figure1

Figure2

Figure 2. Pasupuleti, et al.

Figure3

Figure 3. Pasupuleti, et al.

Figure4

Figure 4. Pasupuleti, et al.

Figure5

Figure6

Figure7

Figure 7. Pasupuleti, et al.

Figure8

Figure 8. Pasupuleti, et al.

Figure9

Figure10

Graphical Abstract:

Fatty acid alkyl esters (FAAEs) were analyzed by gas chromatography (GC) hyphenated with a tandem differential mobility spectrometer (DMS). FAAEs in blends of biodiesel with petrodiesel were selectively analyzed using atmospheric pressure chemical ionization reactions in the detector with mobility characterization of gas ions. Limits of detection were ~ 0.5 ng injected with average 2.98% relative standard deviation (RSD) for peak area precision, ≤ 1.3 % RSD for retention time precision, and ≤ 9.2 % RSD for compensation voltage. The feedstock of biodiesel/petrodiesel blends was classified using principal component analysis (PCA) and hierarchical cluster analysis (HCA) with and without Fisher ratio feature selection supervision. Unsupervised cluster analysis captured 52.72% of variance in a single PC while supervised analysis captured 71.64% of variance. PCA and HCA models built using a training set correctly clustered independent test set samples based on biodiesel feedstock. Fast isolation and detection of methyl linolenate (C18:3 me), a specially regulated fuel component, was achieved in under 10 seconds with a 1 meter long capillary column using dual ion filtering in the tandem DMS.