SUPPORTING INFORMATION

Rapid Analysis and Time Interval Deconvolution for Comprehensive Fuel Compound Group Classification and Speciation using Gas Chromatography – Vacuum Ultraviolet Spectroscopy

Phillip Walsh[1], Manuel Garbalena[2], and Kevin A. Schug[3]*

[1] VUV Analytics, Inc., Cedar Park, TX

[2] McKee Laboratory, Valero Energy Corporation, Sunray, TX

[3] Department of Chemistry and Biochemistry, The University of Texas at Arlington, Arlington TX

*Correspondence to: 700 Planetarium Pl.; Box 19065; Arlington TX 76019-0065; (ph) 817-272-3541; (fax) 817-272-3808; (email) kschug@uta.edu

Contents

TID Algorithm – Tiered Library Search

This section describes the details of the tiered library search portion of the TID algorithm, performed once for each TI in the chromatographic dataset.

Given a total absorbance spectrum and a reduced library set for a TI, a model was constructed based on Equation 2 that consisted of a single compound. A series of linear fits were performed against the total absorbance spectrum of the time slice, where each fit consisted of the one-component model with one of the reduced library reference spectra substituted for analyte 1. A fit metric, in this case the standard χ^2 statistic, was used to rank the quality of each of the one-component fits, and the analyte yielding the best fit metric retained, along with the value of the optimized f_1. Next, the reduced list was searched again, but this time applying Equation 2 to a two-component system. The reference spectra for each possible pair of compounds in the reduced library list were substituted for $A_{1,ref}$ and $A_{2,ref}$, and the fit performed to find optimized f_1 and f_2. Again, each fit was characterized by the χ^2 fit metric, and the fit yielding the best χ^2 value was compared to the best fit from the one-component search. If the best fit from the two-component search was significantly better than the best fit from the one-component search, the two component result was retained, along with the optimized f_1 and f_2 values from the best two-component fit. Otherwise, the two-component result was rejected and the best one-component result retained. Here, a threshold was applied on the percent change (e.g., 40%) of the best one- and two-component χ^2 values. The procedure was repeated, but for all possible triplets of compounds from the reduced library list. Again, if the best three component model was significantly better than the best two-component fit, the result was updated with the best three component model along with optimized f_1, f_2, and f_3 values. The tiered search could in principle be continued to consider any number of components, but for this work, a practical limit of three possible coelutions in any given time interval was imposed. This limit has to be met in reality by the separation conditions and also by

optimizing the size of the time intervals. In a more general case, instead of setting a maximum on the number of coelutions, the algorithm could simply be run to convergence and ended once a stopping criteria is met. A flow diagram illustrating the tiered search for the case of maximum three coelutions is given in Figure S1.

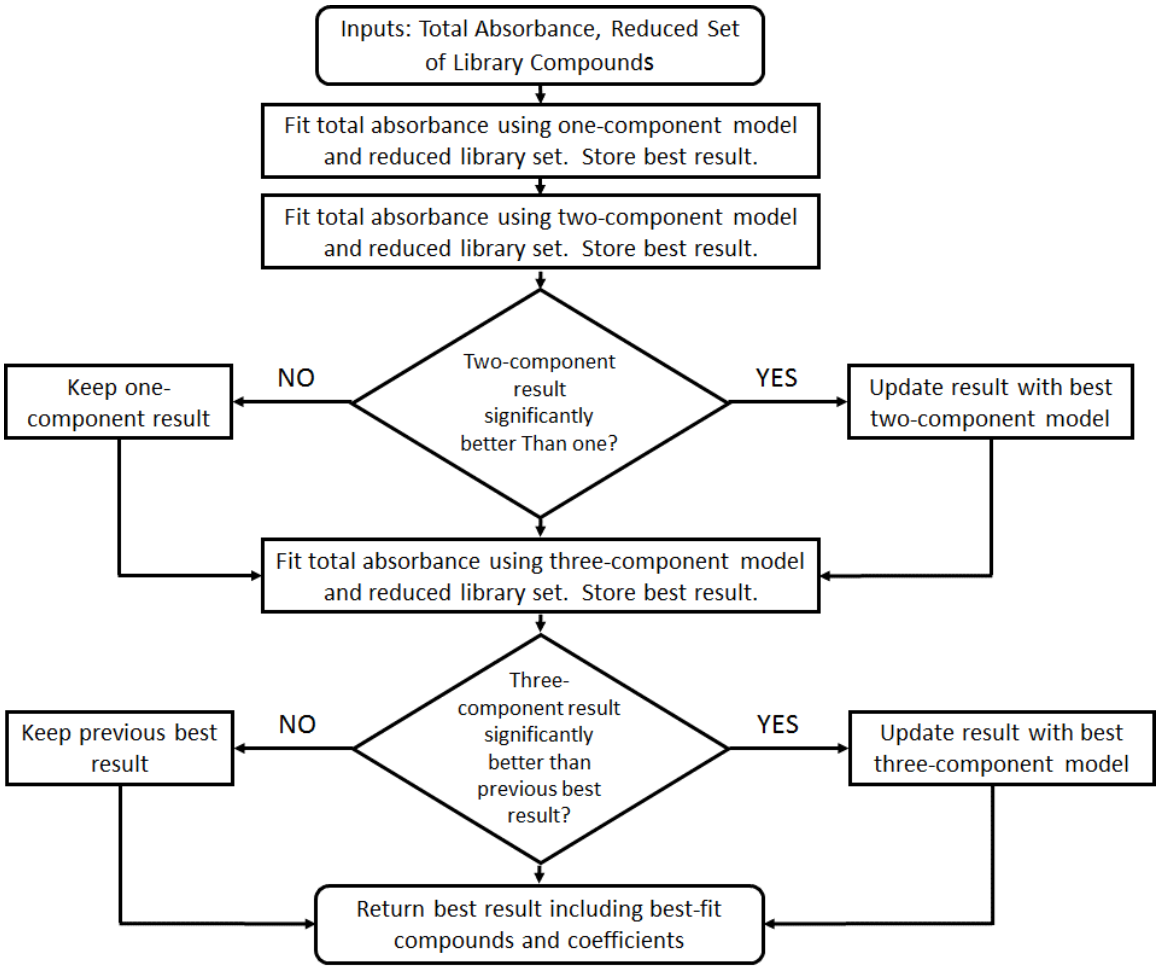
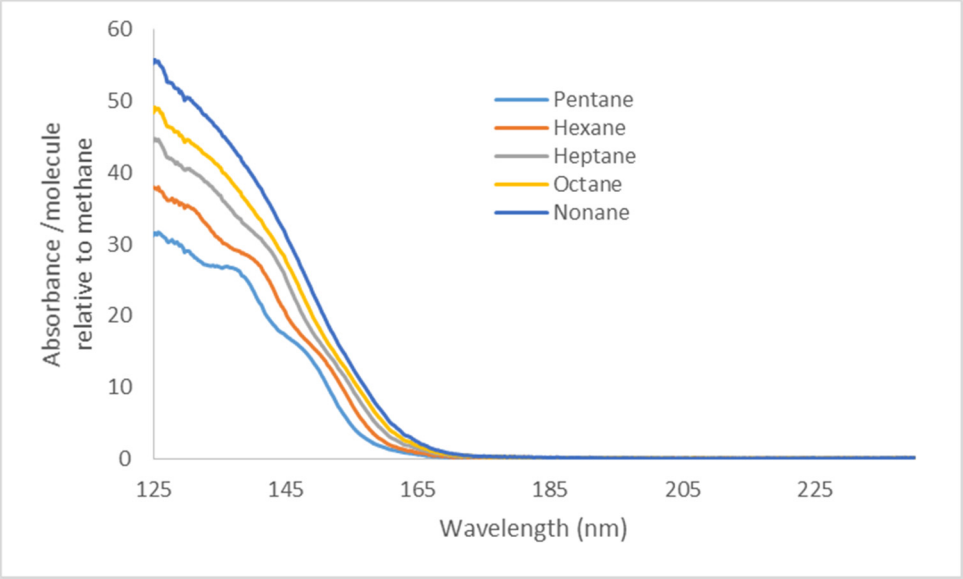


Figure S1. Flow diagram of tiered search portion of TID algorithm.

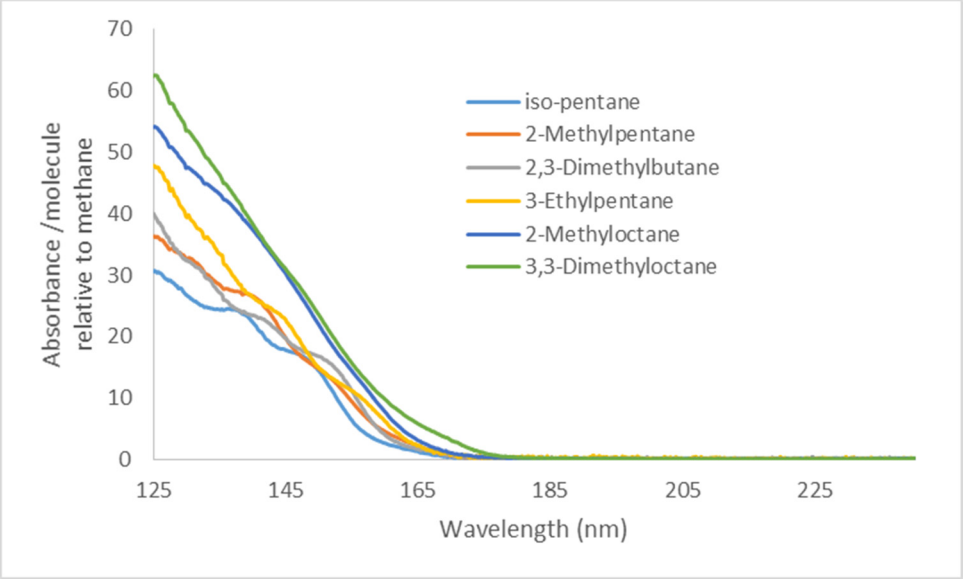
Representative VUV Spectra for PIONA Class Analytes

Figures S2 A. – J. are some class-segregated VUV spectra obtained for various PIONA analytes. The wavelength-dependent shapes of the absorbance spectra are determined by the compounds' identities. There are various ways to scale the spectra in order to compare the relative responses. Presented below are two such comparisons, one on an absorbance per molecule basis (Figures S2 A. – E.) and one on absorbance per unit mass (Figures S2 F. – J.). In both cases, the responses are normalized by the corresponding quantity for methane, averaged over the 125-240nm wavelength region. Thus, for Figures S2 A. – E., each response is normalized by the average 125-240nm methane absorbance per molecule, and for Figures S2 F. – G, by the average 125-240nm methane absorbance per unit mass. Note that while the absorbance per molecule may increase within a given class as the molecules get heavier, the absorbance for a given mass of compound tends to be relatively uniform.

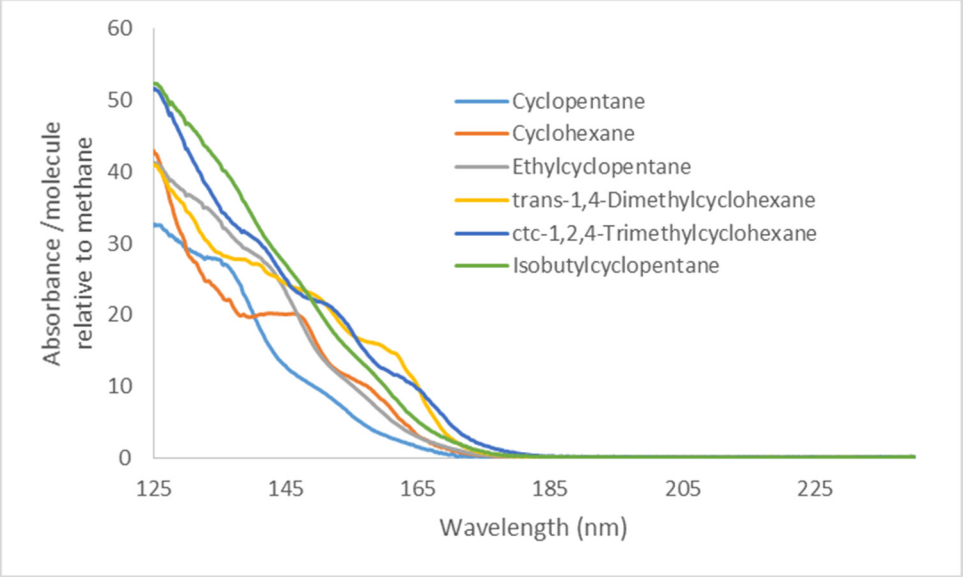
A. Paraffins



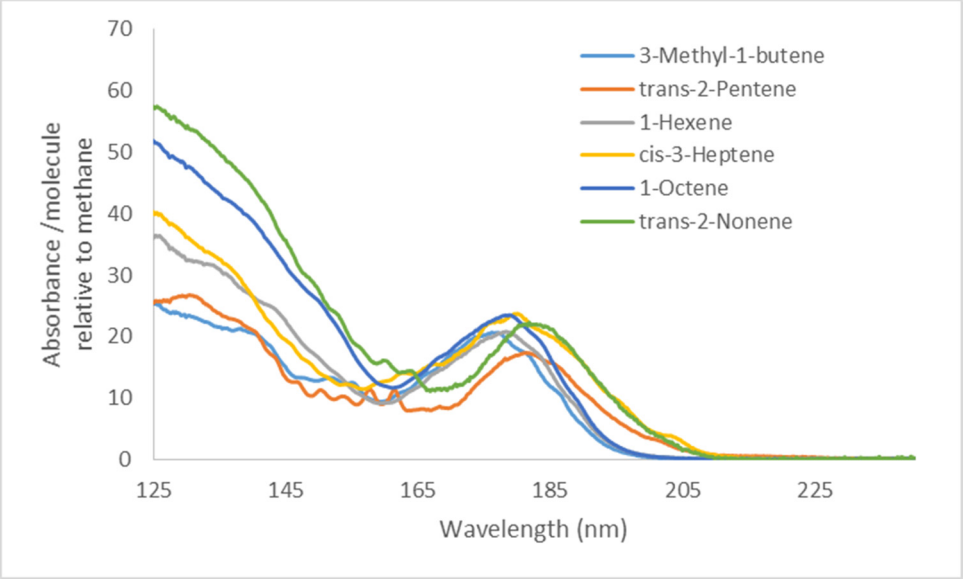
B. Isoparaffins



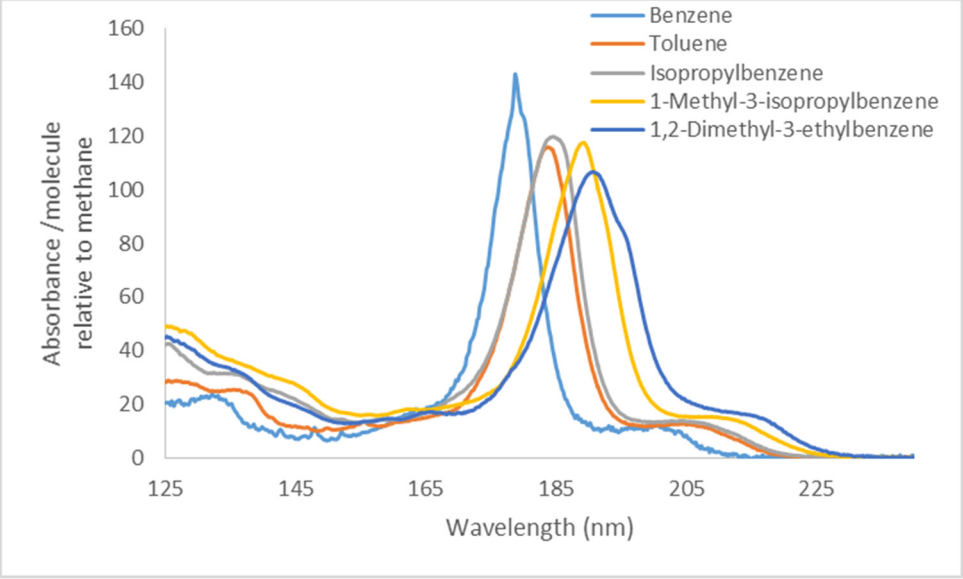
C. Naphthenes



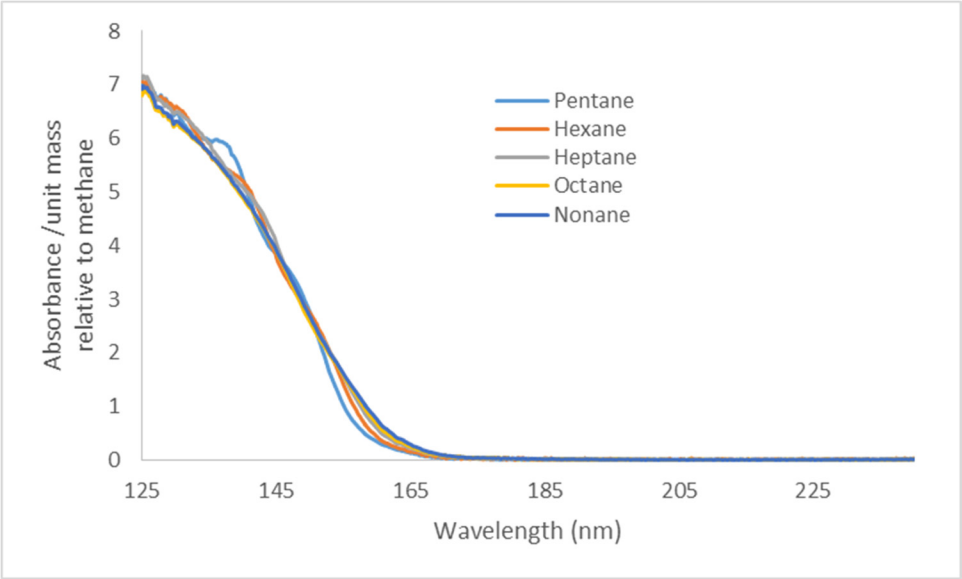
D. Olefins



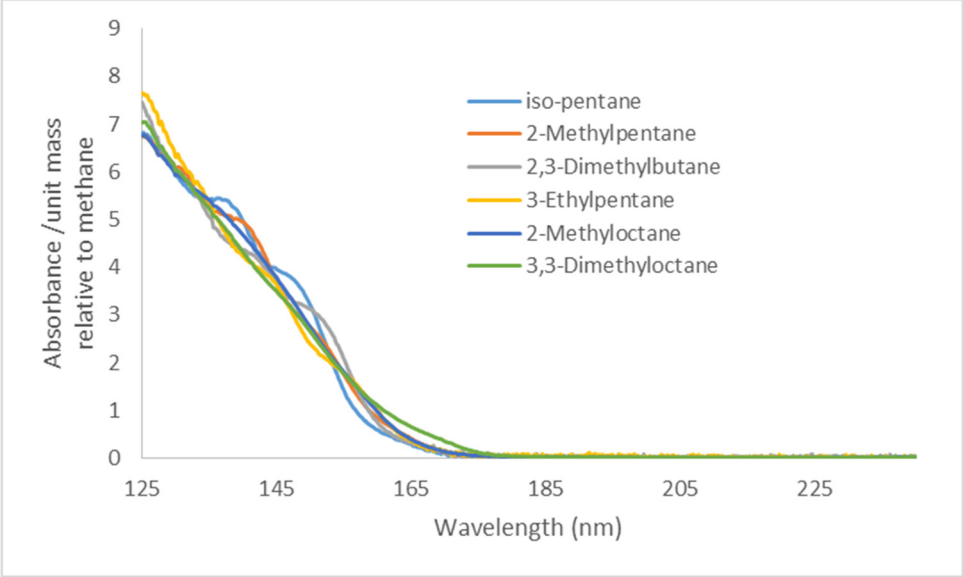
E. Aromatics



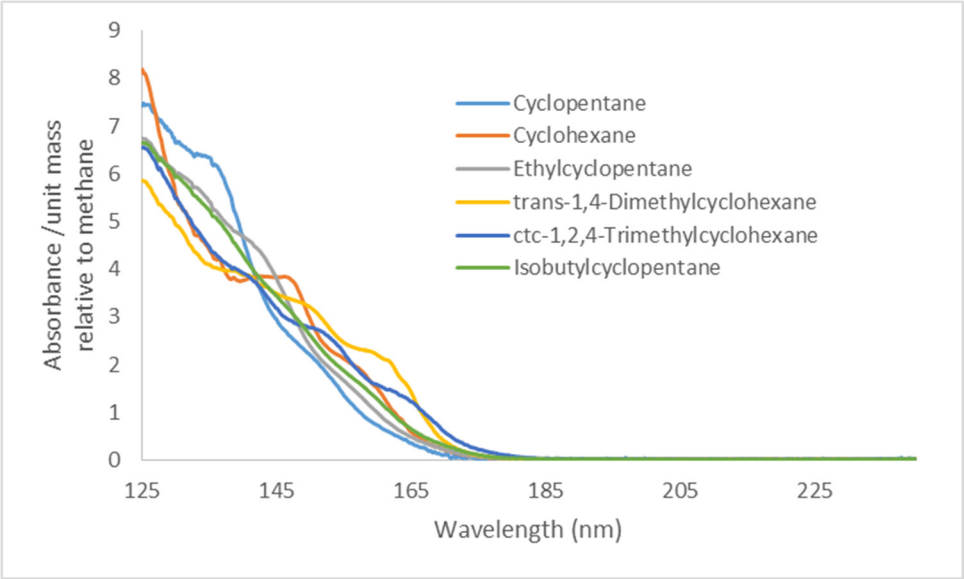
F. Paraffins



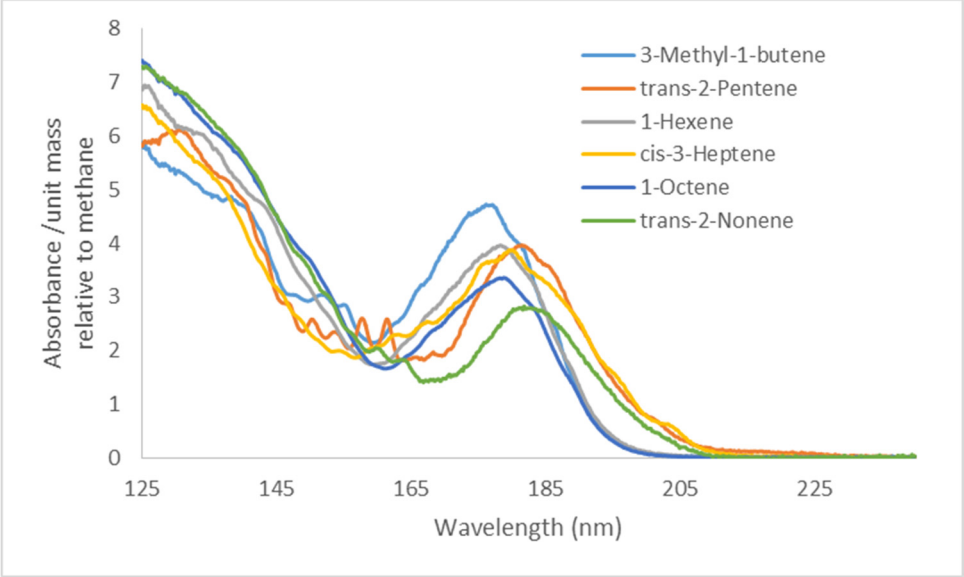
G. Isoparaffins



H. Naphthenes



I. Olefins



J. Aromatics

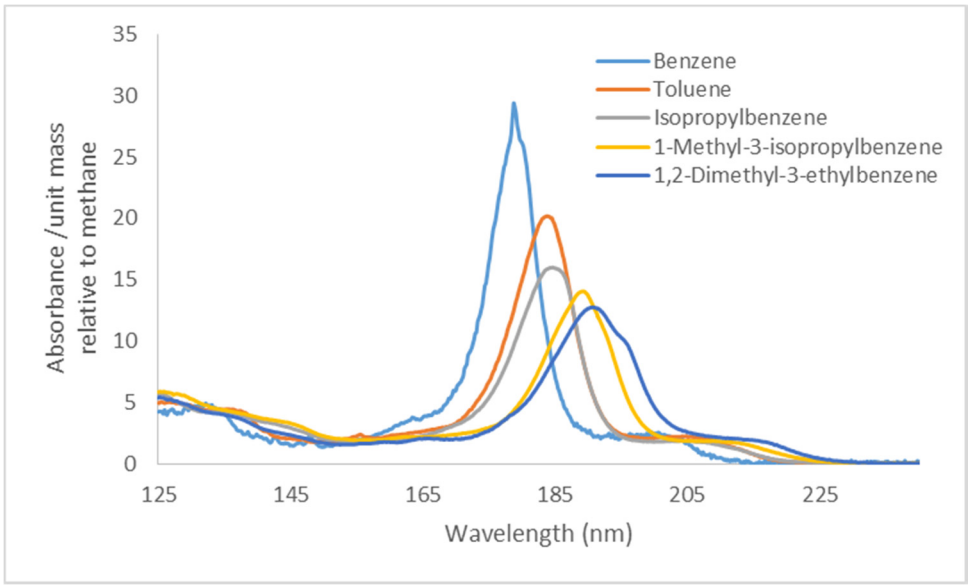


Figure S2. Representative VUV spectra for different PIONA compound classes.

Class-Based Relative Response Factors for GC-VUV PIONA Analysis

Starting values for PIONA class-based RRFs were determined by applying Equation 4 to the 139-component PIONA standard. To do this, the TID algorithm was applied to a chromatographic measurement of the 139-component combined PIONA standard up to just prior to applying Equation 3. The total response areas for each of the PIONA classes and class total mass % values from the standard certification were used to derive class RRFs relative to the paraffins class by substituting bulk class quantities in place of "analytes" 1 and 2 in Equation 4. This procedure determined the class-based RRFs of the five PIONA classes relative to each other, but left the paraffins RRF undetermined.

VUV cross sections for the linear alkanes are available and were used to determine individual paraffin RRFs via Equation 5, relative to methane. A plot showing the carbon number dependence of the paraffins RRFs is given in Figure S3 A. The data show that the areal responses per unit mass for the C5+ linear alkanes are very similar. Therefore, a class-based paraffins RRF was derived from the average C5 – C13 theoretical paraffin RRFs. This class-based response factor was used to finish the calibration of the other four PIONA classes and also to represent the paraffins as a class in subsequent gasoline analyses.

For samples having high percentages of C1 – C4 linear alkanes, Figure S3 A. suggests that compound-specific RRFs should be considered for the lower carbon numbers. This amounts to a carbon number breakdown of the paraffins RRFs. A carbon number breakdown for the other classes can be performed, if desired, using suitable standards along with Equation 4. In this case, the RRFs can be determined for several compounds within each class and the RRFs for compounds having the same carbon number averaged. Alternately, Equation 4 can be applied to carbon number groupings of compounds; M_1 and A_1 refer to averages over groups of compounds belonging to the same class and carbon number, similarly for M_2 and A_2.

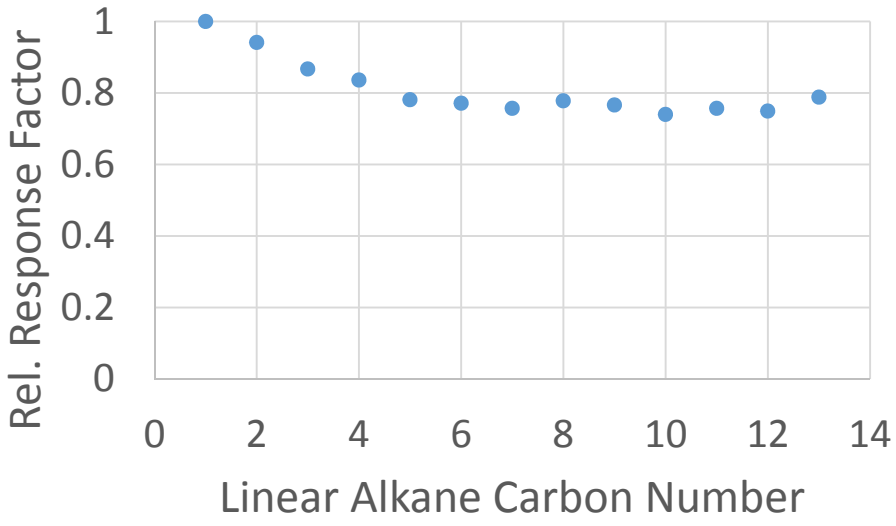
Further motivation for using simpler bulk class-based RRFs, at least with respect to gasoline-range analyses, is shown in Figures S3 B. – E. In each figure, one of the class-specific background samples was manually analyzed to determine the percent of the total peak area represented by each individual compound. These relative peak areas are plotted against the reported mass % values provided with the PIONA standard documentation. In the handful of cases where the long 100 m separation was not able to baseline-resolve the reference components, the overlapping peaks were combined, as were the standard mass % values of the offending compounds. Consideration of Equation 3 shows that if the RRFs for all of the compounds within a given class were identical, each plot would be perfectly linear, with no scatter and a slope of 1. The fact that each plot is nearly so is an indication that the within-class RRFs are very similar. Repetitions under controlled conditions was not done, but the dataset was obtained on three separate occasions using similar conditions, and the data points in the plots are the averages of the three runs and relative peak area analyses. Error was estimated from the range of the three measurements, represented by an error bar on each data point. There was no accuracy information in the standard documentation, so horizontal error bars have not been included and the stated mass % values are taken as given.

Note that each plot in Figure S3 compares data only from compounds of the same class. The RRFs for different classes can be very different. Also note that RRFs depend on the wavelength region considered. The uniformity of within-class RRFs is strongest when considering a response that includes a wide wavelength region, in this case the 125-240 nm region. The scatter apparent in the plots increases when the response covers narrower regions, and in an extreme case where the response is determined by only a single wavelength value, there is no apparent similarity at all.

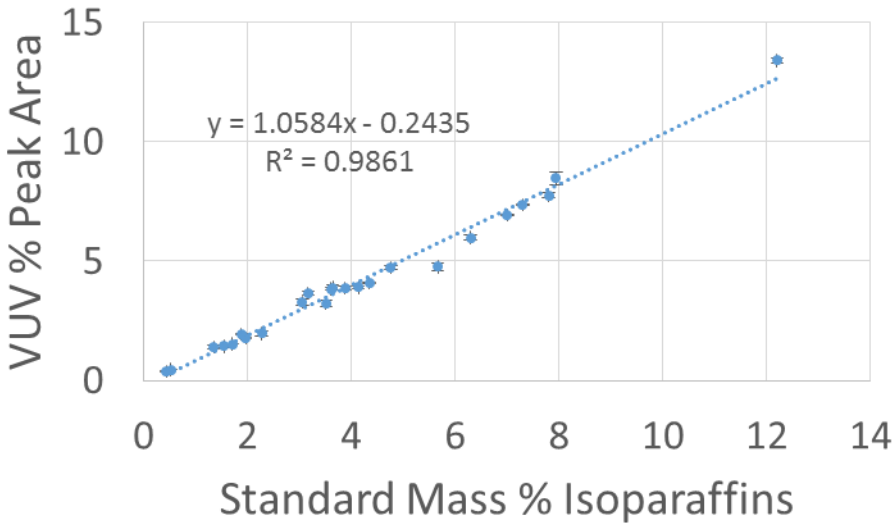
The similarities of within-class RRFs, along with similar within-class absorbance spectra have major implications for petroleum fuel analyses. These two characteristics together significantly simplify combined classification and speciation analyses of such complex samples. These properties are

expected to have an even greater impact when extending similar analyses to heavier fuels, such as jet fuel and diesel. In these cases it will likely be necessary to increase the number of distinct classes considered, as well as to incorporate explicit carbon number dependence of class RRFs.

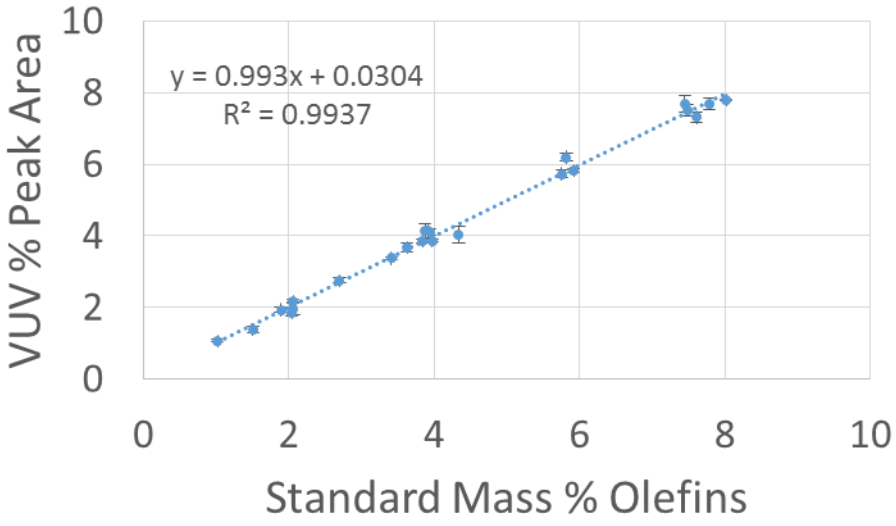
A.



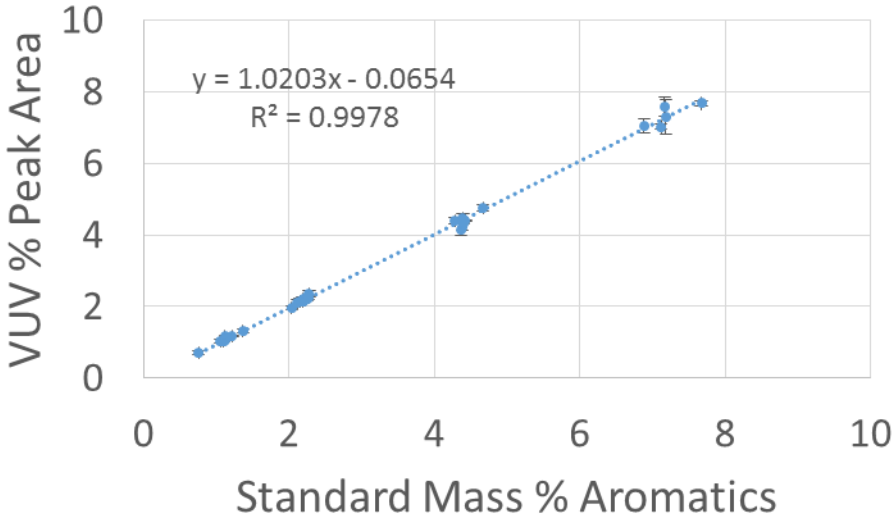
B.



C.


y = 0.993x + 0.0304
R² = 0.9937
VUV % Peak Area
Standard Mass % Olefins
0
2
4
6
8
10

D.


y = 1.0203x - 0.0654
R² = 0.9978
VUV % Peak Area
Standard Mass % Aromatics
0
2
4
6
8
10

E.

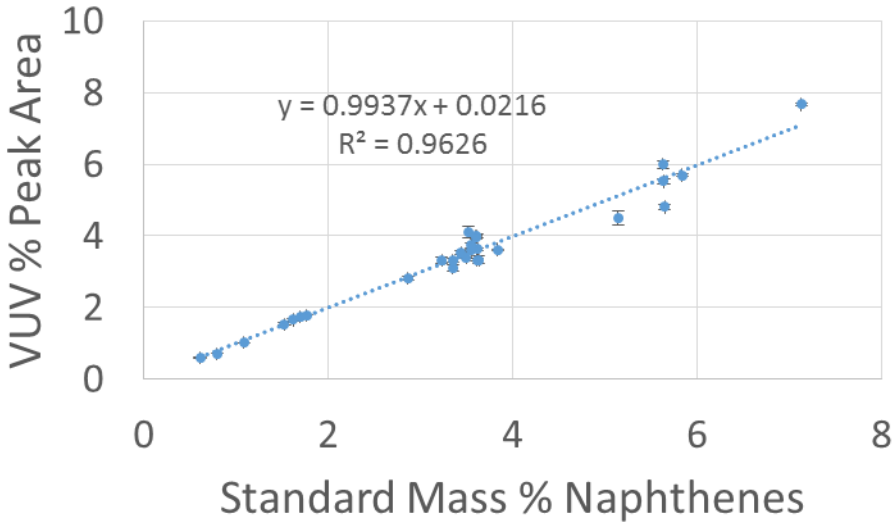
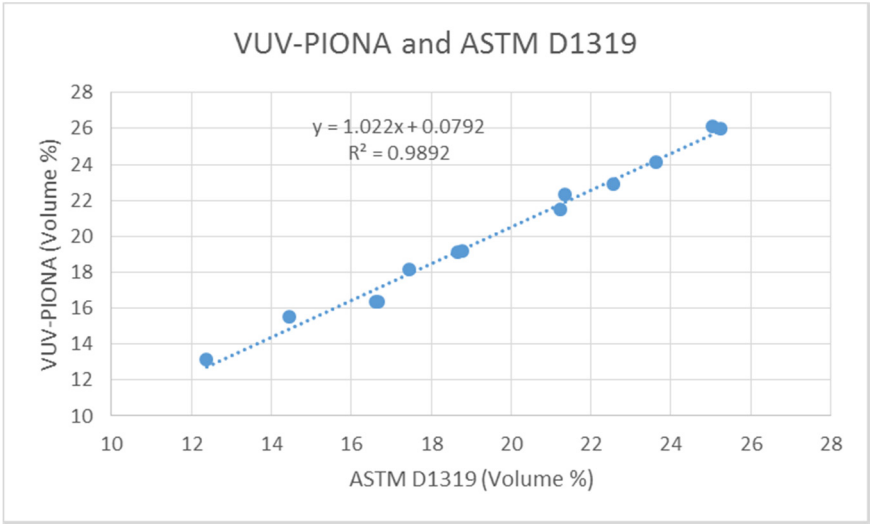


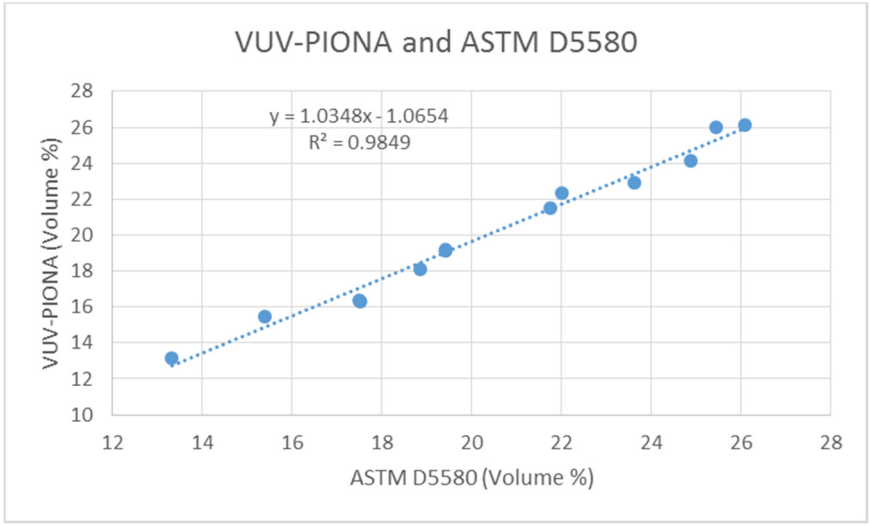
Figure S3. (A.) illustrates the variation of VUV relative response factors (relative to methane) with carbon number for the paraffins compound class. (B. – E.) show correlations between the standard mass % reported in the PIONA standard documentation with measured GC-VUV relative peak area for different analytes, segregated by compound class: (B.) Isoparaffins; (C.) olefins; (D.) aromatics; and (E.) naphthenes.

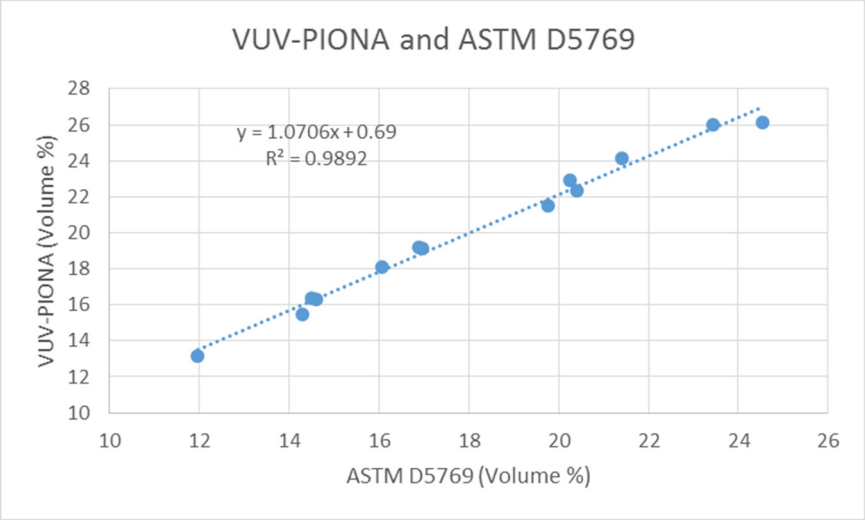
Correlation of GC-VUV versus Various ASTM Test Results for Gasoline Proficiency Samples

Figures S4 A.- F. show the results of individual comparisons of the correlations between GC-VUV results and ASTM test results for different PIONA classifications and component speciations. These data correspond to the data shown in Figure 5 of the main text.

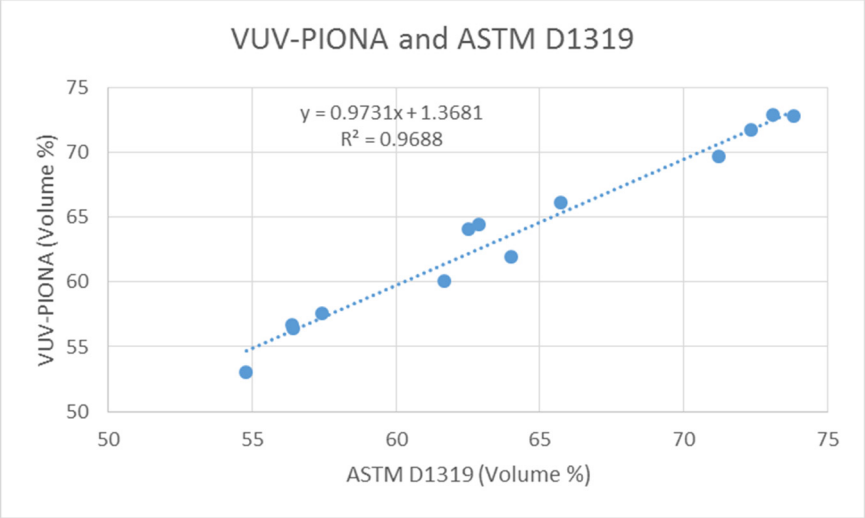
A. Total Aromatics



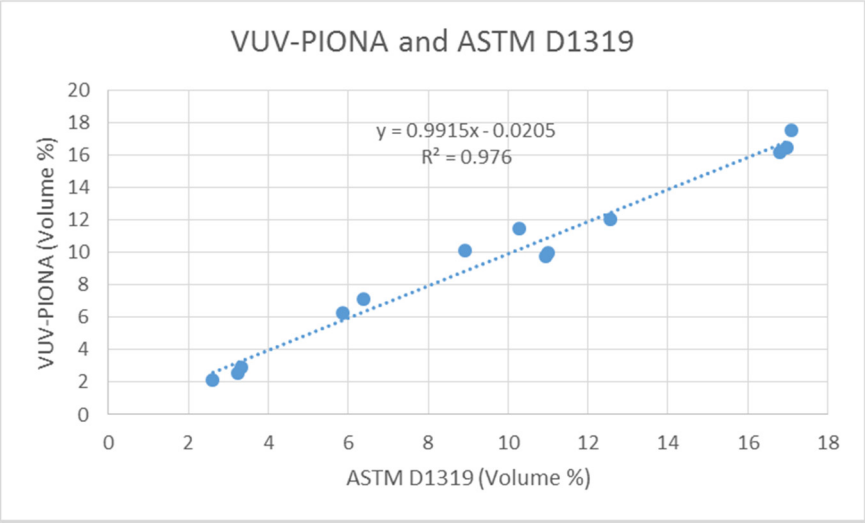


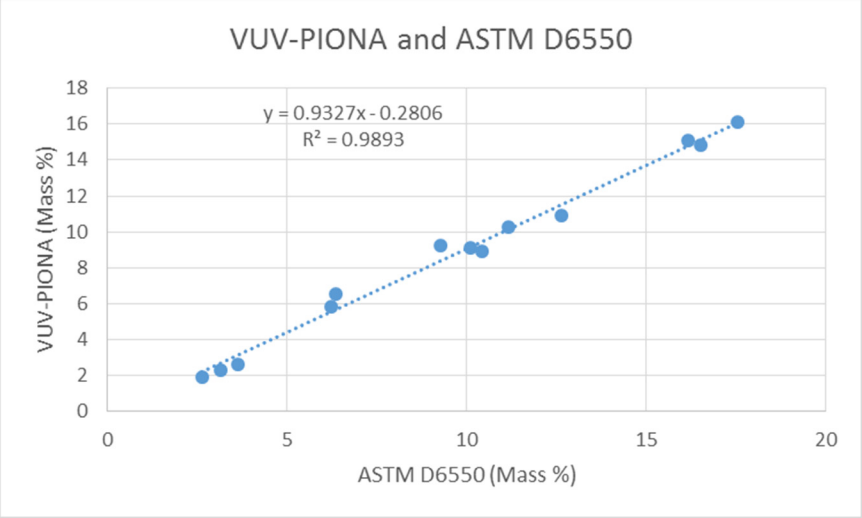


B. Total Saturates

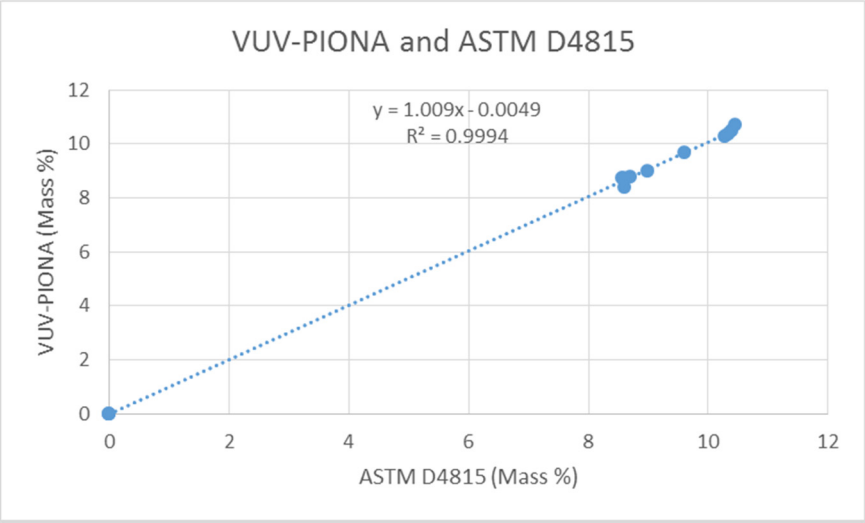


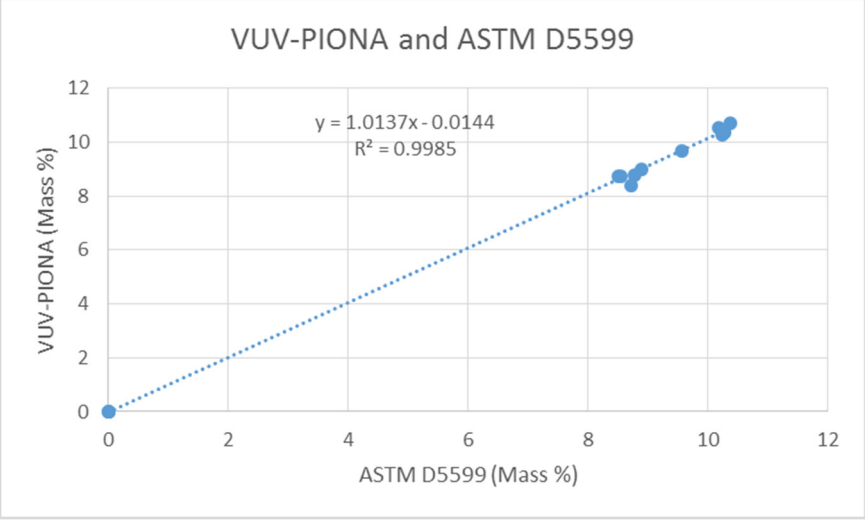
C. Total Olefins

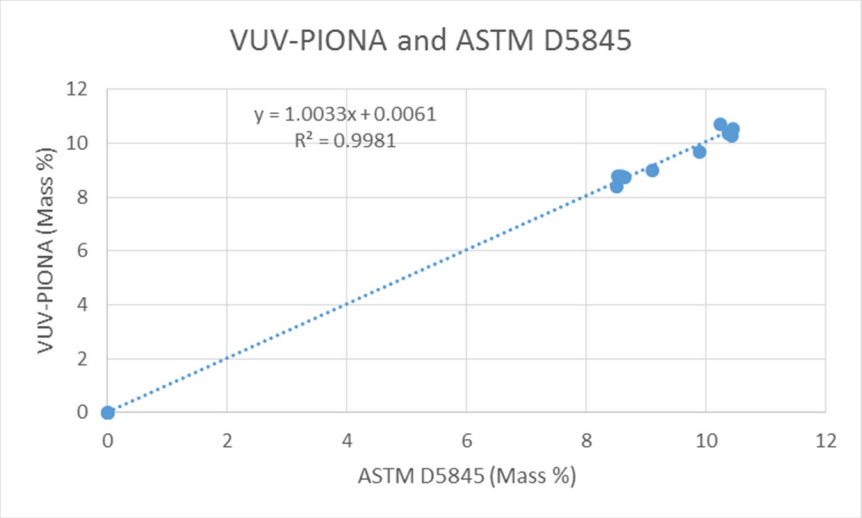




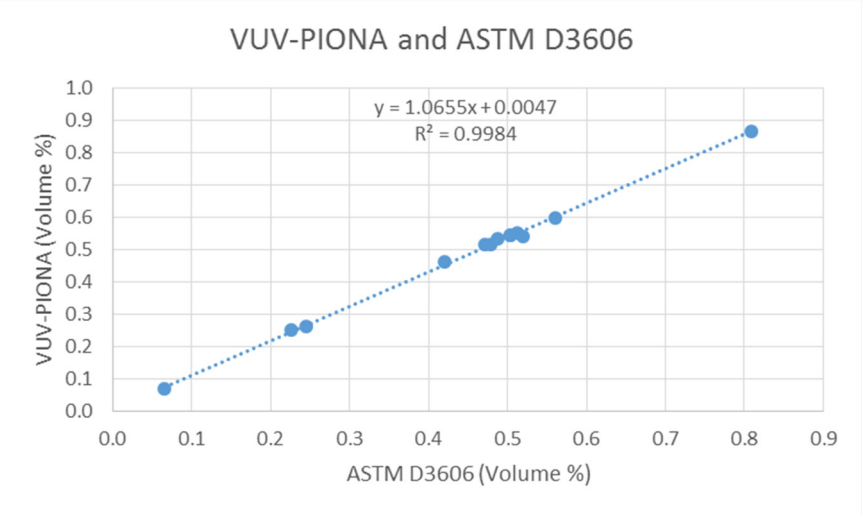
D. Ethanol

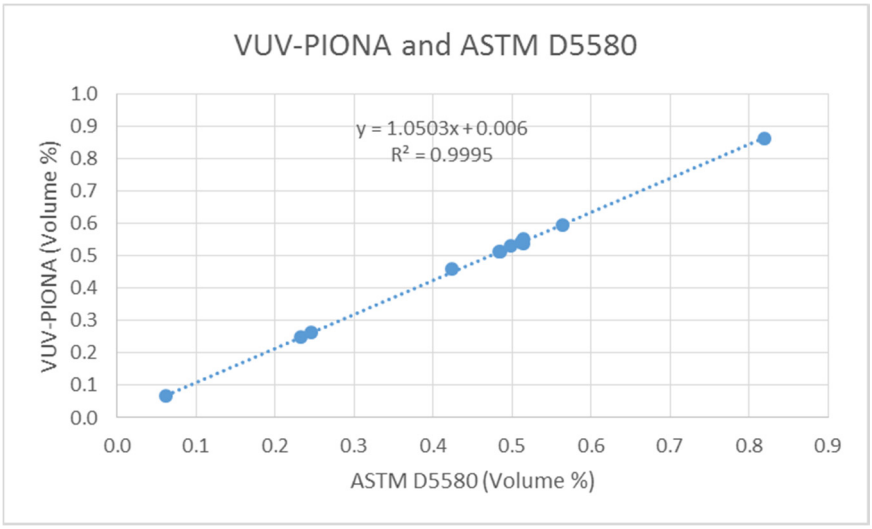

VUV-PIONA and ASTM D4815
y = 1.009x - 0.0049
R² = 0.9994
VUV-PIONA (Mass %)
ASTM D4815 (Mass %)
0
2
4
6
8
10
12

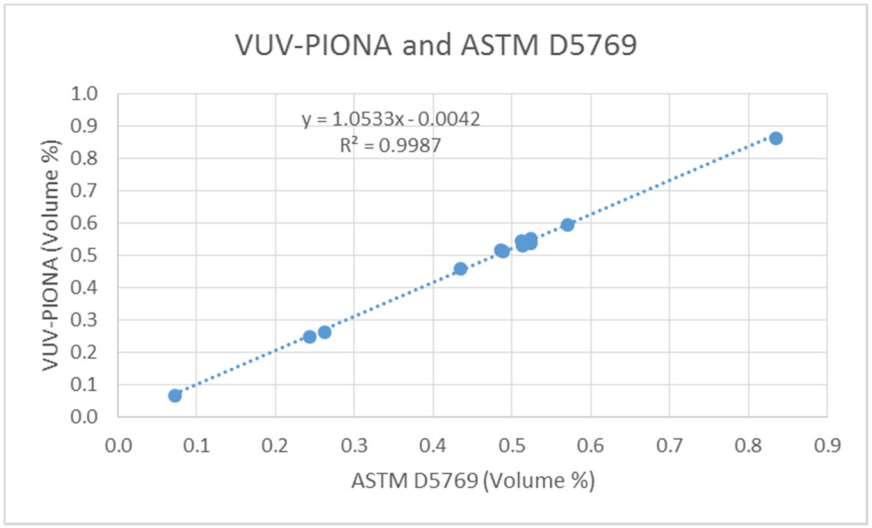

VUV-PIONA and ASTM D5599
y = 1.0137x - 0.0144
R² = 0.9985
VUV-PIONA (Mass %)
ASTM D5599 (Mass %)
0
2
4
6
8
10
12


VUV-PIONA and ASTM D5845
y = 1.0033x + 0.0061
R² = 0.9981
VUV-PIONA (Mass %)
ASTM D5845 (Mass %)
0
2
4
6
8
10
12

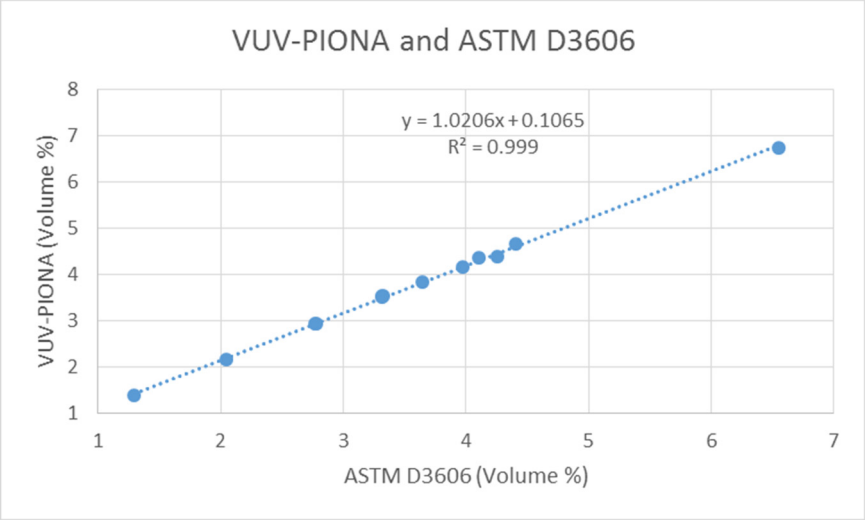
E. Benzene

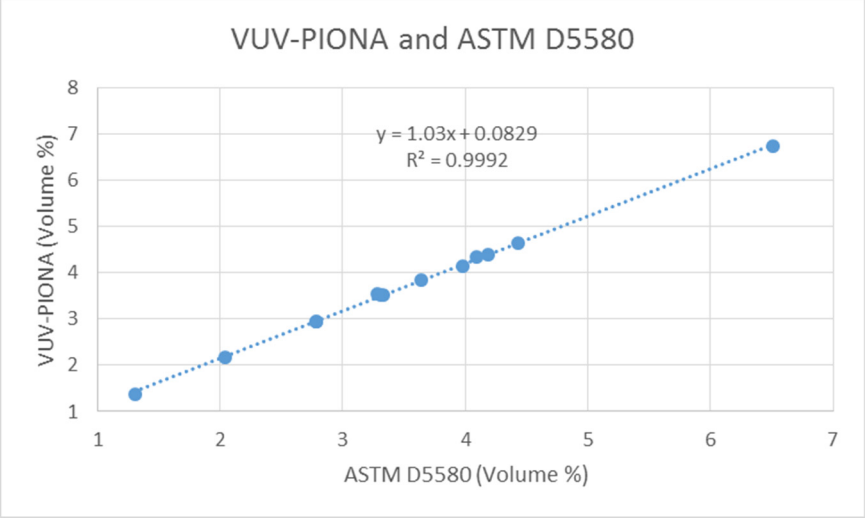

VUV-PIONA and ASTM D3606
y = 1.0655x + 0.0047
R² = 0.9984
VUV-PIONA (Volume %)
ASTM D3606 (Volume %)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


VUV-PIONA and ASTM D5580
y = 1.0503x + 0.006
R² = 0.9995
VUV-PIONA (Volume %)
ASTM D5580 (Volume %)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


VUV-PIONA and ASTM D5769
y = 1.0533x - 0.0042
R² = 0.9987
VUV-PIONA (Volume %)
ASTM D5769 (Volume %)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

F. Toluene


VUV-PIONA and ASTM D3606
y = 1.0206x + 0.1065
R² = 0.999
VUV-PIONA (Volume %)
ASTM D3606 (Volume %)
1
2
3
4
5
6
7
8


VUV-PIONA and ASTM D5580
y = 1.03x + 0.0829
R² = 0.9992
VUV-PIONA (Volume %)
ASTM D5580 (Volume %)
1
2
3
4
5
6
7
8

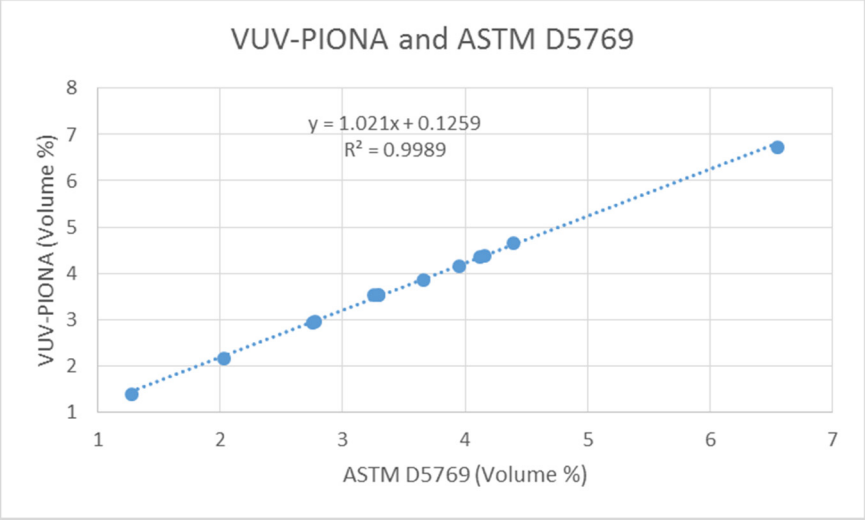
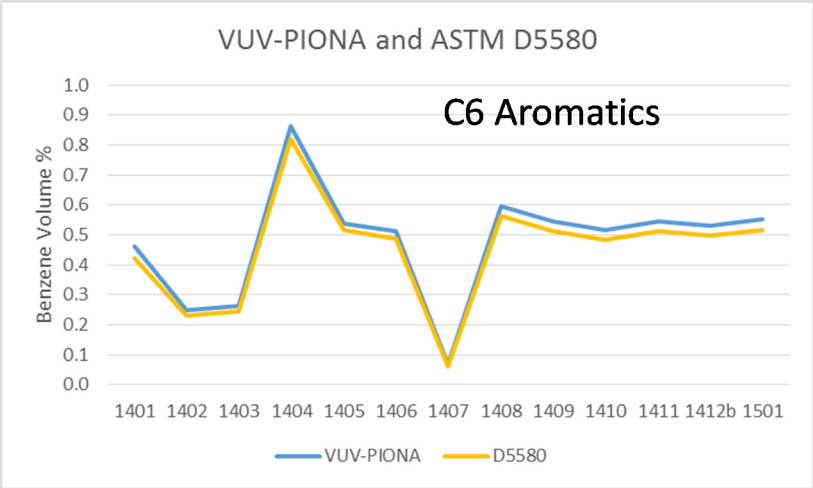
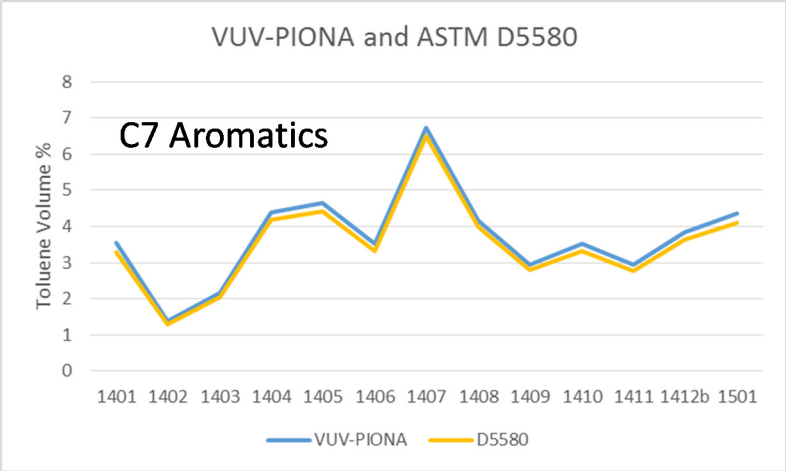


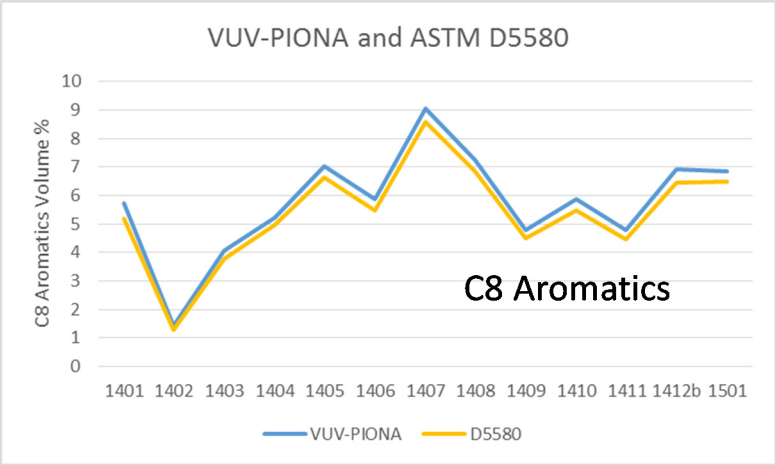
Figure S4. Correlations between GC-VUV and various ASTM results for a series of thirteen gasoline proficiency samples.

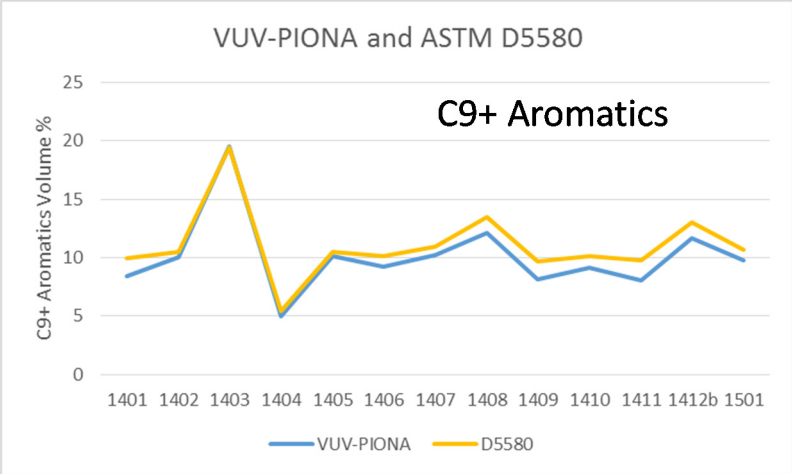
Comparison of GC-VUV and ASTM D5580 for Benzene, Toluene, C8 Aromatics, C9+ Aromatics, and Total Aromatics for Gasoline Proficiency Samples

Figure S5: ASTM D5580 proficiency tests reported an additional category consisting of C9 and heavier aromatics. Since the reported D5580 results also included benzene, toluene, and total aromatics, the contribution from C8 aromatics was inferred by subtracting benzene, toluene, and C9+ aromatics from the total aromatics, as determined by ASTM D5580.









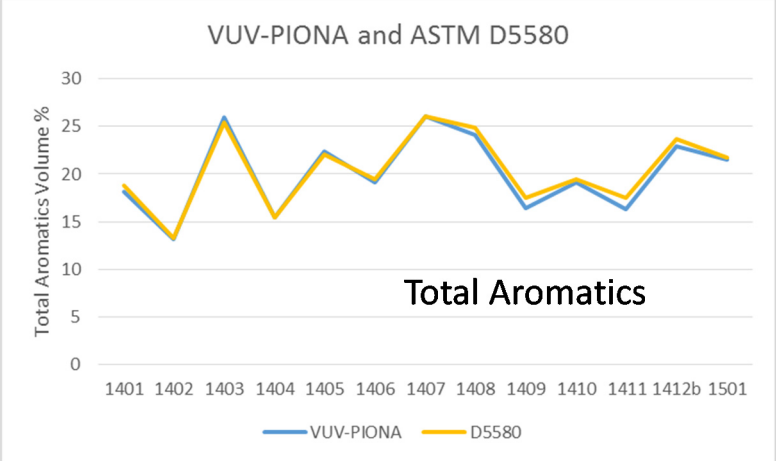


Figure S5. Comparison of GC-VUV and ASTM D5580 measurements of benzene, toluene, C8, C9+, and total aromatics for a series of 13 gasoline proficiency samples.

Carbon Number PIONA Distributions for Gasoline Proficiency Standards using GC-VUV TID

Using results from the GC-VUV separation and the subsequent TID analysis procedure, carbon number distributions of each PIONA class in terms of mass % for all thirteen gasoline proficiency samples tested are shown in Figure S6.

Figure S6. Carbon number distributions of PIONA classes in ASTM gasoline proficiency samples