CRC Report No. RW-121a/CM-138-22a

Statistical Analysis of Nationwide PMI Fuel Survey

Final Report

October 2025

COORDINATING RESEARCH COUNCIL, INC.
1 CONCOURSE PARKWAY • SUITE 800 • ATLANTA, GA 30328

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries with participation from other industries, companies, and governmental bodies on research programs of mutual interest. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The five main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance; heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars); and sustainable mobility (e.g., decarbonization). CRC's function is to provide the mechanism for joint research conducted by industries that will help in determining the optimum combination of products. CRC's work is limited to research that is mutually beneficial to the industries involved. The final results of the research conducted by, or under the auspices of, CRC are available to the public.

FINAL REPORT

Statistical Analysis of Nationwide PMI Fuel Survey

CRC Project

RW-121a / CM-138-22a

Prepared by:

Andrey Khlystov and S. Kent Hoekman

Desert Research Institute
Reno, NV 89512

August 8, 2025

Table of Contents

1. Executive Summary 8

2. Introduction and Background 12

2.1. Particulate Matter Emissions from Gasoline Vehicles 12

2.2. Particulate Matter Index 12

2.3. Methods for PMI Determination 13

3. Project Objectives 14

4. Methods 15

4.1. Data sets 15

4.2. Data Harmonization 15

4.2.1. Chemical composition data harmonization 15

4.2.2. Distillation profile harmonization 16

4.3. Data Exploration and Statistical Analysis 17

5. Results 18

5.1. Summer dataset 18

5.1.1. Overview of the summer dataset 18

5.1.2. Chemical composition data intercomparison 18

5.1.3. Summer distillation profile intercomparison 24

5.1.4. Group contributions to weight and PMI 26

5.1.5. Comparison of summer gasolines across grades and PADD regions 28

5.1.6. Comparison of PMI across grades and PADD for summer dataset 33

5.1.7. Contributions of generic compounds to weight and PMI in summer fuels 35

5.1.8. Summer distillation profiles 36

5.1.9. Hypothetical effect of distillation cuts on summer fuel PMI 37

5.2. Winter dataset 41

5.2.1. Overview of the winter dataset 41

5.2.2. Chemical composition data intercomparison for the winter dataset 41

5.2.3. Winter distillation profile intercomparison 46

5.2.4. Group contributions to weight and PMI 46

5.2.5. Comparison of winter gasolines across grades and PADD regions 48

5.2.6. Comparison of PMI across grades and PADD for winter dataset 52

5.2.7. Contributions of generic compounds to weight and PMI in winter fuels 55

5.2.8. Winter distillation profiles 55

5.2.9. Hypothetical effect of distillation cuts on winter fuel PMI 56

5.3. Comparison of summer and winter datasets 58

5.3.1. Comparison of fuel composition and PMI between the two seasons 58

5.3.1. Comparison of distillation profiles between the two seasons 64

5.3.2. Comparison with the 2010 Honda study 65

6. Conclusions 67

7. References 70

List of figures

Figure 1. A map of PADD regions. 8
Figure 2. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the summer dataset. The descriptive statistics of the ratios are also provided. 9
Figure 3. Histograms of fuel PMI for summer and winter datasets (all grades and PADD regions) 11
Figure 4. Cumulative distributions of PMI values for summer and winter datasets of this project and that of the Honda 2010 study. The PMI values in this figure do not include contribution of compounds for which PMI factors are not known. 11
Figure 5. Comparison of total weight % of individual PIONAX groups reported by the D6730 and D8369 methods for the summer data set. The dotted line shows the 1:1 relationship, while the solid blue line shows the regression line with the regression equation and the corresponding R^2 of the fit given in the lower right corner of each subplot. 20
Figure 6. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the summer dataset. The descriptive statistics of the ratios are also provided. 21
Figure 7. Per compound weight comparison within individual PIONAX groups reported by the D6730 and D8369 methods for the summer dataset. Compounds with the same carbon number are indicated by the same color (carbon numbers are given in the legends). The dotted line shows the 1:1 relationship 22
Figure 8. Comparison of concentrations of two potentially coeluting pairs (toluene and 2,3,3-trimethylpentane, and benzene and 1-methylcyclopentene) reported by the D6730 and D8369 methods for the summer dataset. 23
Figure 8. Left: Comparison of PMIs reported by D6730 (FID) and D8369 (VUV) methods for the summer dataset ("Not interpolated" – PMI data as reported, "Interpolated" – including interpolated PMI data). The broken black line represents a 1:1 line. Right: Histogram of FID/VUV ratios (the insert provides descriptive statistics of the ratios; "pmi" – PMI data as reported, "pmi_int" – including interpolated PMI data). 24
Figure 9. Comparison of the average distillation profiles measured with the D86 and D7096 methods and derived from the D6730 and D8369 data for the summer dataset. 25
Figure 10. Violin plots and boxplots of per-sample differences in reported temperatures required to evaporate a certain fraction of fuel for the summer dataset. 26
Figure 11. Average contributions of the PIONAX groups to fuel weight composition and PMI for the summer dataset. 27
Figure 12. Compound group contributions to weight and PMI for the summer dataset based on the FID (D6730) data. 28
Figure 13. Contribution of aromatic compounds to weight and PMI for the D6730 summer dataset. PMI values were calculated including interpolated values for generic compounds. 28
Figure 14. Average composition of the three gasoline grades per PIONAX group and carbon number based on D6730 summer data (RU: regular unleaded, PU: premium unleaded). 29
Figure 15. Statistical comparison of chemical composition of the three fuel grades using D6730 summer data (RU: regular unleaded, PU: premium unleaded). Three stars denote $p < 0.001$, two stars: $p < 0.01$, one star: $p < 0.05$, and "ns" for p values above 0.05 (not significant at 5% confidence). 30
Figure 16. Statistical comparison of chemical composition of summer premium unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote $p < 0.001$, two stars: $p < 0.01$, one star: $p < 0.05$, and "ns" for p values above 0.05 (not significant at 5% confidence). 31
Figure 17. Statistical comparison of chemical composition of summer regular unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote $p < 0.001$, two stars: $p < 0.01$, one star: $p < 0.05$, and "ns" for p values above 0.05 (not significant at 5% confidence). 32

Figure 18. Statistical comparison of chemical composition of E15 gasoline in PADD regions 1,2,3 broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence). 33
Figure 19. Statistical comparison of PMI of different grades of fuel and fuels in different PADD regions for the entire summer dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence). .. 34
Figure 20. Statistical comparison of PMI between different PADD regions for the three fuel grades for the summer dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence). 34
Figure 21. The weight fraction of generic compounds within aromatic ('a'), isoparaffin ('i'), olefin ('o'), and naphthene ('n') groups for the summer D6730 dataset (all grades and PADD). 35
Figure 22. PMI contribution of generic and unidentified compounds to the total fuel PMI based on the D6730 summer dataset. 'a': aromatics, 'i': isoparaffins, 'o': olefins, 'n': naphthenes, 'u': unidentified. 36
Figure 23. Violin and boxplots of distillation profiles (temperatures required to evaporate a certain fraction of fuel) for summer dataset (alld grades and PADD)..... 37
Figure 24. Temperature at which 95% (T95), 97% (T97), and 99% (T99) of fuel is evaporated vs. fuel PMI for the summer dataset for the four measurement methods. The D8369 subplot uses PMI values derived using the D8369 method; the other plots use PMI values derived using the D6730 method. 39
Figure 25. Estimated PMI after a distillation cut (95%: T95, 97%: T97, and 99%: T99) as a function of the original PMI for the summer D6730 dataset..... 40
Figure 26. Comparison of total weight % of individual PIONAX groups reported by the D6730 and D8369 methods for the winter dataset. The dotted line shows the 1:1 relationship, while the solid blue line shows the regression line with the regression equation and the corresponding R^2 of the fit given in the lower right corner of each subplot..... 42
Figure 27. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the winter dataset. The descriptive statistics of the ratios are also provided. 43
Figure 28. Per compound comparison within individual PIONAX groups reported by the D6730 and D8369 methods for the winter dataset. Compounds with the same carbon number are indicated by the same color (carbon numbers are given in the legends). The dotted line shows the 1:1 relationship..... 44
Figure 29. Left: Comparison of PMIs reported by D6730 (FID) and D8369 (VUV) methods for the winter dataset ("Not interpolated" – PMI data as reported, "Interpolated" – including interpolated PMI data). A broken black line represents 1:1 relationship. Right: Histogram of FID/VUV ratios (the insert provides descriptive statistics of the ratios; "pmi" – PMI data as reported, "pmi_int" – including interpolated PMI data)..... 45
Figure 30. Comparison of the average distillation profiles measured with the D86 and D7096 methods and derived from the D6730 and D8369 data for the winter dataset. 45
Figure 31. Violine plots and boxplots of per-sample differences in reported temperatures required to evaporate a certain fraction of fuel for the winter dataset. 46
Figure 32. Average contributions of the PIONAX groups to fuel weight composition and PMI for the winter dataset. 47
Figure 33. Compound group contributions to weight and PMI for the winter dataset based on the FID (D6730) data..... 48
Figure 34. Contribution of aromatic compounds to weight and PMI for the D6730 winter dataset. PMI values were calculated including interpolated values for generic compounds. 48
Figure 35. Average composition of the three gasoline grades per PIONAX group and carbon number based on D6730 winter data (RU: regular unleaded, PU: premium unleaded). 49
Figure 36. Statistical comparison of chemical composition of the three fuel grades using D6730 winter data (RU: regular unleaded, PU: premium unleaded). Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence)..... 50

Figure 37. Statistical comparison of chemical composition of winter premium unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).....51
Figure 38. Statistical comparison of chemical composition of winter regular unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).....52
Figure 39. Statistical comparison of PMI of different grades of fuel and fuels in different PADD regions for the entire winter dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence)...53
Figure 40. Statistical comparison of PMI between different PADD regions for two fuel grades for the winter dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).....53
Figure 41. The weight fraction of generic compounds within aromatic ('a'), isoparaffin ('i'), olefin ('o'), and naphthene ('n') groups for the winter D6730 dataset (all grades and PADD regions).....54
Figure 42. PMI contribution of generic and unidentified compounds to the total fuel PMI based on the D6730 winter dataset (all grades and PADD regions). 'a': aromatics, 'i': isoparaffins, 'o': olefins, 'n': naphthenes, 'u': unidentified.....54
Figure 43. Violin and boxplots of distillation profiles (temperatures required to evaporate a certain fraction of fuel) for winter dataset.....55
Figure 44. Temperature at which 95% (T95), 97% (T97), and 99% (T99) of fuel is evaporated vs. fuel PMI for the winter dataset for the four measurement methods. The D8369 subplot uses PMI values derived using the D8369 method; the other plots use PMI values derived using the D6730 method. 56
Figure 45. Estimated PMI after a distillation cut (95%: T95, 97%: T97, and 99%: T99) as a function of the original PMI for the winter D6730 dataset. 57
Figure 46. Comparison of weight contributions of PIONAX groups between complete summer and winter datasets. 59
Figure 47. Comparison of weight contributions of PIONAX groups between premium unleaded fuels of summer and winter datasets. 60
Figure 48. Comparison of weight contributions of PIONAX groups between regular unleaded fuels of summer and winter datasets. 61
Figure 49. Comparison of weight contributions of PIONAX groups between E15 fuels of summer and winter datasets. 62
Figure 50. Comparison of PMI contributions of PIONAX groups between complete summer and winter datasets. 63
Figure 51. Histograms of fuel PMI for summer and winter datasets. 64
Figure 52. Comparison of average distillation profiles between summer and winter datasets for different measurement methods.....65
Figure 53. Cumulative distributions of PMI values for summer and winter datasets of this project and that of the Honda 2010 study. 67

List of Tables

TABLE 1. THE NUMBER OF GASOLINE SAMPLES IN THE SUMMER DATASET PER FUEL GRADE AND PADD REGION. 18
TABLE 2. MAXIMUM RETENTION INDEXES FOR THE ENTIRE SUMMER DATASET AS REPORTED FOR D6730 AND D8369 ANALYSES. 18
TABLE 3. AVERAGE CONTRIBUTIONS OF PIONAX GROUPS TO WEIGHT AND PMI FOR SUMMER DATASET. "PMI INTERPOLATED" INCLUDES INTERPOLATED PMI VALUES FOR GENERIC AND UNKNOWN COMPOUNDS. 27
TABLE 4. NUMBER OF SAMPLES FOR EACH OF THE MEASUREMENT METHOD PER FUEL GRADE AND PADD. 41
TABLE 5. MAXIMUM RETENTION INDEXES FOR THE ENTIRE WINTER DATASET AS REPORTED FOR D6730 AND D8369 ANALYSES. 42
TABLE 6. AVERAGE CONTRIBUTIONS OF PIONAX GROUPS TO WEIGHT AND PMI FOR WINTER DATASET. "PMI INTERPOLATED" INCLUDES INTERPOLATED PMI VALUES FOR GENERIC AND UNKNOWN COMPOUNDS. 47
TABLE 7. DESCRIPTIVE STATISTICS OF PMI VALUES IN WINTER AND SUMMER DATASETS OF THIS STUDY AND THOSE ESTIMATED FOR THE US FUELS OF HONDA 2010 STUDY. "PMI": PMI VALUES EXCLUDING INTERPOLATED VALUES; 'PMI_INT" PMI VALUES INCLUDING INTERPOLATED VALUES 66

List of Acronyms and Abbreviations

ASTM	American Society for Testing and Materials
CRC	Coordinating Research Council
DBE	double bond equivalent
DHA	detailed hydrocarbon analysis
E15	gasoline/ethanol blend containing nominally 15 volume percent ethanol
FID	flame ionization detector
GC	gas chromatography
GDI	gasoline direct injection
LDGV	light-duty gasoline vehicles
PADD	Petroleum Administration for Defense Districts
PEI	particulate evaluation index
PIONAX	paraffins, isoparaffins, olefins, naphthenes, aromatics, oxygenates
PM	particulate matter
PME	Index of PM emissions
PMI	PM Index
(R+M)/2	anti-knock index as the average of research and motor octane numbers
RVP	Reid vapor pressure
RI	retention index
SIDI	spark ignition direct injection
SimDis	simulated distillation by wide-bore capillary gas chromatography
SSI	Separation Systems, Inc.
VUV	vacuum ultraviolet

1. Executive Summary

The 2010 Honda SAE paper (2010-01-2115) established Particulate Matter Index (PMI) as a key metric for predicting gasoline fuel's tendency to form particulates, calculated from detailed hydrocarbon analysis, double bond equivalents, and vapor pressure data. Since the publication of that study 15 years ago, the market fuel PMI distributions could have changed due to changes in fuel formulation driven by evolving regulations, refinery processes, and ethanol blending requirements, which warrants their re-evaluation. To this end, this project analyzed U.S. fuel property data from two fuel surveys conducted by the Alliance for Automotive Innovation: the summer 2022 fuel survey (CRC project RW-121) and the winter 2023 fuel survey (CRC project RW-121-2). The summer survey contained 169 fuel samples, including 100 regular unleaded gasoline, 44 premium unleaded gasoline, and 25 E15 gasoline fuels. The winter survey contained 85 fuels, including 53 regular unleaded gasoline, 28 premium unleaded gasoline, and 4 E15 gasoline fuels. Samples came from all five PADD regions (Figure 1). Fuels from both surveys were analyzed for their chemical composition using two methods: (1) the ASTM D6730 method that uses gas chromatography – flame ionization detector (GC-FID) and (2) the ASTM D8369 method that uses gas chromatography - vacuum ultraviolet (VUV) spectroscopy (GC-VUV). In addition to the chemical analysis, the distillation profiles of fuel samples were measured using ASTM D7096 (SimDis) and ASTM D86 (boiling range by atmospheric distillation).

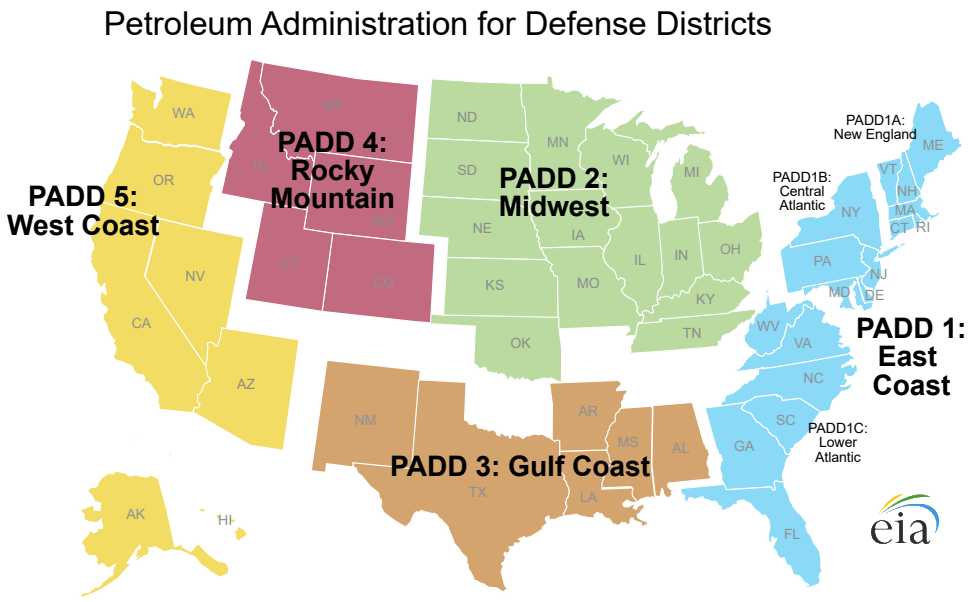

Figure 1. A map of PADD regions.

The chemical composition data were harmonized to the same naming convention, allowing calculation of the PM index (PMI) from the detailed chemical composition data, as well as estimation of the fuel distillation profiles. The distillation profiles measured with the ASTM D86 and D7096 methods were also harmonized to the same temperature and / or distillation steps. The data were then analyzed using various statistical techniques to gain insight into the relative performance of the analytical methods and any differences in the chemical composition and PMI of fuels across the seasons, fuel grades, and PADD regions.

Both the D6730 and D8369 methods demonstrated a remarkable performance in identifying compounds or at least assigning a compound class to a chromatographic peak, which reduces the uncertainty in PMI determination. The VUV method reported no unidentified compounds, while the FID method reported only 0.2 weight percent in summer and 0.1 weight percent in winter as unknown compounds, with negligible contributions to the overall PMI. The contribution of generic compounds (i.e., those for which only the compound class was identified) is also small. Based on the D6730 data, generics contributed only 1% of all aromatic weight. The highest contribution of generics was observed for naphthenes, where they contributed approximately 10%-13% (winter and summer, respectively) of all weight of that compound group. Their contribution to the PMI was highest for aromatics, being about 5% of the aromatic-related PMI.

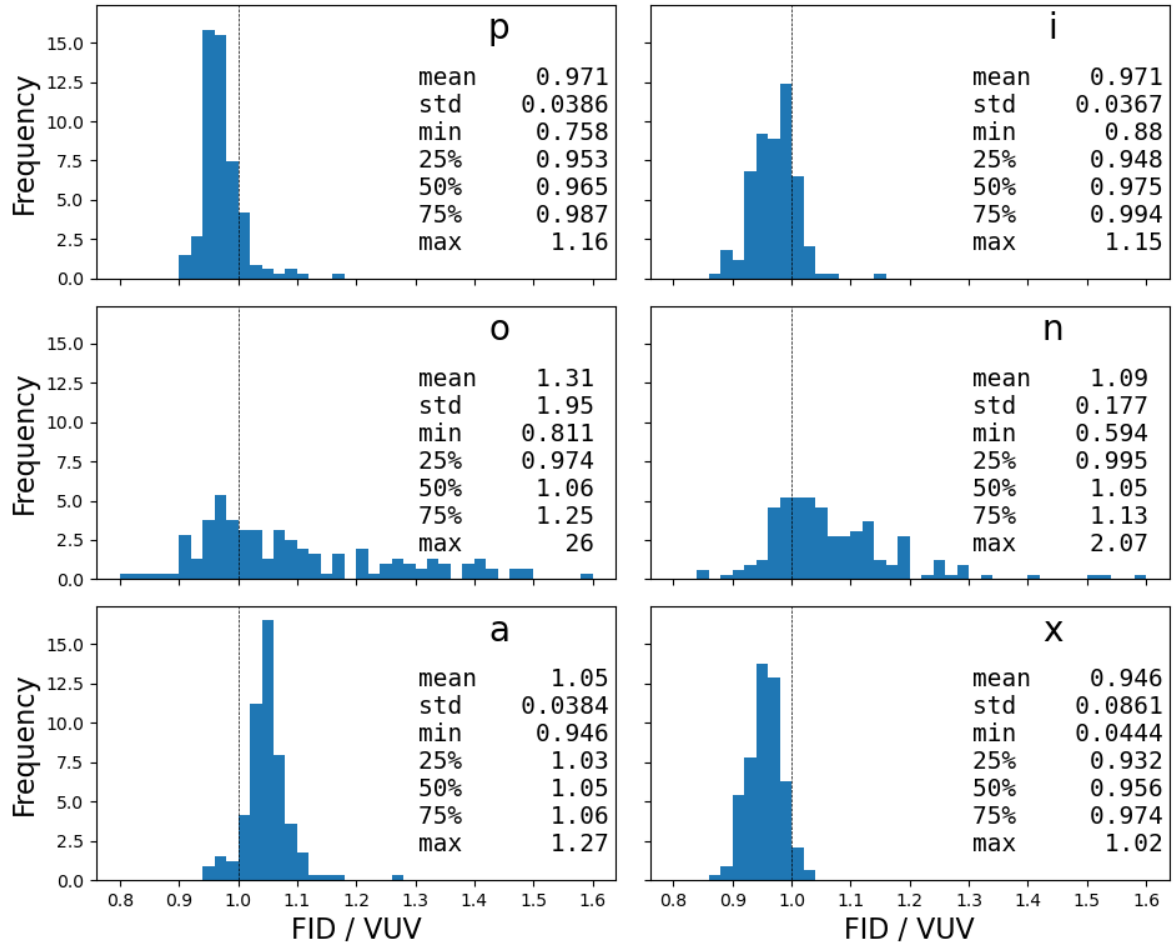

Figure 2. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the summer dataset. The descriptive statistics of the ratios are also provided.

A per-compound comparison of the D6730 and D8369 methods indicates that the two methods correlate well for compounds with weight percent contributions above approximately 1%. The agreement between the two methods below 1 weight % is very poor, with individual values differing by up to two orders of magnitude. The large discrepancy at small concentrations could be a result of the baseline noise interference and/or the ability of the two methods to discriminate between partially co-eluting compounds. Overall, the D8369 results are biased high for paraffins,

isoparaffins, and oxygenates relative to the D6730, while the opposite is observed for olefins, naphthenes, and aromatics (Figure 2). PMI values derived using the D6730 and D8369 methods also correlated well with each other, but the FID method reported consistently higher PMI than the VUV method (by 25% in summer and 28% in winter). The reason for these discrepancies is not clear, as it cannot be explained by the differences in the chromatographic range of the two methods. It should be noted, however, that the D6730 had a better sensitivity, larger chromatographic range, a significantly better chromatographic resolution, and reported approximately 70% more chromatographic peaks within the same RI range.

The distillation profiles also differed among the four methods. Generally, the agreement between the methods was better for the intermediate boiling range, while becoming worse at each end of the distillation profile. The composition-derived distillation profiles agreed well with each other, though with a small bias for most of the range, which increased at the low and high boiling point ranges. The SimDis method disagreed significantly with the other two chromatography-based methods (D6730 and D8369), especially in the lighter half of the distillation range. The reason for this discrepancy is not clear. A positive correlation was observed between the fuel PMI and temperatures at which certain fraction of fuel is evaporated (hypothetical distillation cuts). The correlations, however, are not very strong, with R^2 values ranging between 0.41 and 0.69.

Statistically significant differences in chemical composition were observed across different seasons, fuel grades, and PADD regions. The most apparent differences in chemical composition between the grades, aside from the higher ethanol content of the E15 grade, are the higher isooctane and toluene contents of the premium unleaded gasoline relative to the other two grades. This is consistent with these compounds being used to obtain the higher (R+M)/2 for premium grade gasoline. The main difference between the seasons is in the amounts of light compounds, especially butane, that are more abundant in winter than summer. These differences are consistent with the Reid vapor pressure (RVP) limitations – lower in summer than winter. This is also reflected by the distillation profiles, which indicate less volatile gasoline in summer than winter. Aromatic compounds are the main contributor to the fuel PMI, with the C8-12 aromatics (i.e., aromatics with 8 to 12 carbon atoms) having the most prominent contribution. In summer, the PMI of premium unleaded fuels was statistically significantly lower than the other two grades (the regular and E15). Among PADD regions, fuels in PADD 2 (the Midwest region) had statistically significantly higher PMI values than fuels in the other regions, while the differences between the other regions are not statistically significant. Overall, summer fuels have a slightly higher average PMI (1.54) than winter fuels (1.48) (Figure 3). This difference, however, is not statistically significant at 95% confidence level.

There is a clear and significant reduction in fuel PMIs since the Honda 2010 study (Figure 4). The median point, which was 1.69 in the Honda study, reduced for the current survey to 1.21 in winter and 1.31 in summer for D6730. The corresponding mean values reduced from 1.76 to 1.23 in winter and 1.31 in summer for D6730. Greater reductions are observed for the higher end of the PMI distribution.

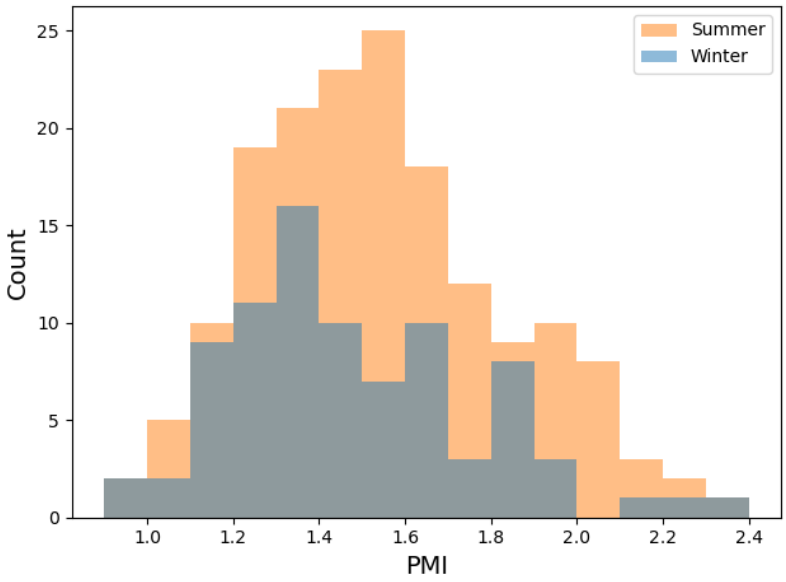

Figure 3. Histograms of fuel PMI for summer and winter datasets (all grades and PADD regions).

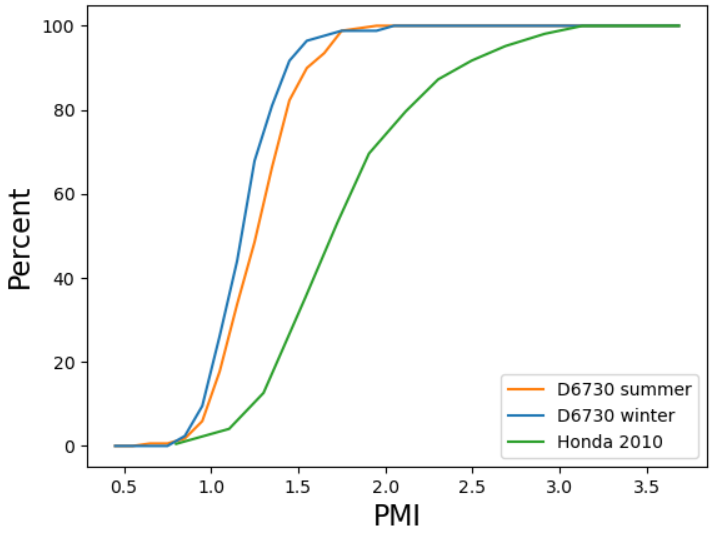

Figure 4. Cumulative distributions of PMI values for summer and winter datasets of this project and that of the Honda 2010 study. The PMI values in this figure do not include contribution of compounds for which PMI factors are not known.

2. Introduction and Background

2.1. Particulate Matter Emissions from Gasoline Vehicles

The continued drive to increase the performance and efficiency of light-duty gasoline vehicles (LDGV) and to reduce their emissions has led to development of modern fuel injection systems such as the gasoline direct injection (GDI) [also called spark ignition direct injection (SIDI)]. While GDI engines have a higher efficiency and specific power output, they produce more particulate matter (PM) than traditional port fuel injection engines.[1] In GDI, liquid fuel is injected directly into the combustion chamber, which can potentially lead to incomplete droplet evaporation and excessive wetting of the piston and combustion chamber walls with liquid fuel, resulting in diffusion burning that causes increased PM emissions.

In addition to improvements in engine combustion chamber and fuel injection system design, PM emissions from LDGV could be reduced by changes in fuel composition, which has been known to influence PM emissions from spark ignition engines. Different chemical constituents within gasoline have different propensities to form PM. In particular, aromatic species are more prone to form soot than are paraffinic species.[2] Aromatics are often thought of as the nucleus around which soot growth occurs. In addition, the boiling point (or vapor pressure) of a compound controls its evaporation processes. Heavier, less volatile fuel species are the last to vaporize, and hence are concentrated in the residual liquid droplets or films that undergo diffusion burning in the combustion chamber. The combination of low volatility and heavy aromatics (such as C_{10+} aromatic compounds) represents the worst situation with respect to PM formation.

2.2. Particulate Matter Index

Several parameters have been developed that link gasoline's properties with its propensity to form PM,[3] including the Honda Particulate Matter Index (PMI)[4] and General Motors Particulate Evaluation Index (PEI).[5] The PMI is a very popular metric that generally provides good predictive accuracy,[6] though with some exceptions.[7] Improvements of the PMI have been proposed, such as the Index for Particulate Matter Emissions (PME).[8] Nonetheless, the PMI metric remains very popular in research on the impact of fuel composition on vehicle emissions. Since the original study by the Honda group also provided a worldwide survey of gasoline fuels available on the market in the summers of 2008 and 2009 using the PMI,[4] this parameter provides a link to that survey, allowing investigation of trends in fuel properties in the past decade.

The PMI is calculated using the following equation:[4]

$$PMI = \sum_{i=1}^{n} \frac{DBE_i + 1}{VP_i} W_i \quad (1)$$

In Equation 1, W_i is the weight percent of gasoline specie i, VP_i is its vapor pressure (in kPa) at a temperature of 443°K (170°C), and DBE_i refers to its "double bond equivalent". For hydrocarbons, DBE is equal to:

$$DBE = 1 + C - 0.5\,H \quad (2)$$

in which *C* and *H* refer to the number of carbon and hydrogen atoms in the specie, respectively. Straight chain paraffins have a DBE value of 0; cyclic paraffins and straight chain single olefinic compounds have DBE values of 1; aromatics have a DBE value of ≥4. Due to their large DBE, aromatics have a strong contribution to a fuel's PMI, especially the heavier compounds due to the inverse relationship of the PMI to the compound vapor pressure.

2.3. Methods for PMI Determination

PMI calculations require highly detailed information about the chemical composition of gasoline, since DBE and VP values are needed for each of the hundreds of species present in the fuel. The original Honda study, as well as most of the follow-up studies on PMI, performed the detailed hydrocarbon analysis (DHA) that uses high resolution gas chromatographic (GC) methods. Two existing ASTM methods used for this purpose are D6729 and D6730.[9-11] In all these cases, capillary GC columns are used to separate volatile fuel constituents, which are then detected and quantified using a flame ionization detector (FID). ASTM D6729 and ASTM D6730 both utilize a 100-meter capillary column of 0.25 mm inside diameter but differ in the use of a short pre-column (used in ASTM D6730) to improve the resolution of a few important pairs of species which otherwise co-elute and cannot be independently measured.

While the degree of species resolution and detection provided by ASTM D6729 and D6730 are sufficient for most purposes, they may not be adequate in certain situations, such as when determining PMI values of gasoline. A particular concern with use of these ASTM methods is their inability to identify and quantify low concentrations of the heaviest (least volatile) constituents in gasoline. These heavy constituents are rich in aromatics, and therefore have relatively high PMI values. Thus, the inability to measure these species, even though they may represent only a small fraction of the total fuel mass, could result in significant underestimation of a gasoline's PMI value.

To improve the resolution of heavy-end compounds, Separation Systems, Inc. (SSI), within a CRC-funded project AVFL-29, modified the standard ASTM D6730 method by including an additional temperature ramp for the GC column. Thus, instead of a final column temperature of 200 °C as specified in D6730, SSI's enhanced method includes an additional, rapid temperature ramp to 300 °C. This has the effect of eluting the highest boiling gasoline compounds as relatively sharp chromatographic peaks that can be integrated and quantified. Additionally, SSI significantly improved the identification of chromatographic peaks, both by use of mass spectrometry (MS) and additional authentic standard compounds. In this way, SSI developed an expanded (and more reliable) list of 822 gasoline constituents.[12] This enhanced method and constituent list were subsequently included in ASTM D6730 as Appendix X1.[10] Gage Products Company has been adding to this component list, most recently with the addition of the 231 components listed in Appendix 1 of this report. This augmented list was used for the D6730 analyses in this project. Henceforth, references to "D6730" in this report will refer to the Gage version of the method.

Another GC-based method for chemical analysis is described in ASTM D8369. This method uses high resolution gas chromatography with Vacuum Ultraviolet Absorption Spectroscopy (GC-VUV) for the determination of individual compounds and compound classes by percent mass or percent volume with a final boiling point as defined by ASTM D86 up to 225 °C. The

separation is accomplished using a 60-meter, nonpolar phase capillary column and a moderately fast temperature ramp. The advantage of using VUV instead of the FID lies in its better specificity, as compound identification using the retention time can be supplemented with the compound VUV spectrum, such that coelutions could be resolved by the detector using vacuum ultraviolet absorbance spectra and deconvolution. The specificity of VUV, however, is significantly lower than that of the MS techniques and thus could lead to less accurate determination of the PMI. A detailed explanation of this method is provided in Appendix 2 of this report.

The main disadvantage of the DHA techniques lies in their complexity and long analysis times (a single chromatographic run on a 100-meter column takes 3-4 hours to complete). To resolve this problem, a simpler and faster (< 15 min) simulated distillation technique (SimDis, ASTM D7096) was proposed as a potential surrogate for DHA based PMI determination. The SimDis uses a non-polar, large-bore, thick film capillary column. The column temperature is raised at a linear rate such that hydrocarbons elute in their boiling point order and are measured by an FID. A linear relationship between the chromatographic retention time and hydrocarbon boiling point is determined using a known mixture of hydrocarbons and volume fractions are determined from chromatographic area slices using theoretical hydrocarbon volume response factors (oxygenated fuels require oxygenate response factors). In this way a boiling point distribution is obtained.

As can be seen from the above description, different PMI determination methods in use today all have their advantages and disadvantages. The PMI formulation requires a highly detailed knowledge of the chemical composition of gasoline, as well as knowledge of vapor pressure of individual compounds. However, even with highly sophisticated chromatographic techniques it is not always possible to identify all compounds, nor are the vapor pressure data available for all compounds. This is especially true for the heavy end of a chromatographic run. The SimDis method, though much simpler and faster than the DHA techniques, by default cannot provide the chromatographic resolution of the DHA methods. All these factors could lead to discrepancies in predicting PM emissions with the DHA and SimDis methods.

3. Project Objectives

The main objectives of the project are to:

1. Perform statistical analysis of data from CRC RW-121 and RW-121-2 projects including PMI based on DHA (per ASTM D6730), SimDis (per ASTM D7096), boiling range by atmospheric distillation (per ASTM D86), and VUV spectroscopy (per ASTM D8369).

2. Identify correlations to methods and/or fuel properties, data gaps and sources of differences between the methods, and determine trends in regional fuel survey data.

3. Summarize and formalize results of Objectives 1 and 2 in two written reports (one for the summer survey RW-121 and one for the winter survey RW-121-2, which will also include a comparison to the summer survey).

4. Methods

4.1. Data sets

This project analyzed data from two fuel surveys: (1) the summer fuel survey (CRC project RW-121) during which 169 fuel samples from July 2022 were analyzed, and (2) the winter fuel survey (CRC project RW-121-2) with 85 fuels from January 2023 analyzed. The fuel samples were obtained from the "Alliance for Automative Innovation North American Fuel Survey©".[16] The Alliance for Automotive Innovation North American Fuel Survey data reflect a single sample "snapshot" of market fuel properties from retail stations sampled in various cities. The number of stations varied from city to city, and cities and stations can vary from survey to survey. In the survey, the cities and stations sampled are not selected to meet statistical criteria, or on the basis of market share.

Fuels from both surveys were analyzed for their chemical composition and for their distillation profiles. For chemical analysis, two methods were used: ASTM D6730 using gas chromatography – flame ionization detector (GC-FID) performed by Gage Products Company and ASTM D8369 using gas chromatography - vacuum ultraviolet (VUV) spectroscopy (GC-VUV) performed by VUV Analytics. The distillation profiles were measured using ASTM D7096 (SimDis) performed by Southwest Research Institute (SWRI) and ASTM D86 (boiling range by atmospheric distillation). The D86 data were obtained from the "Alliance for Automative Innovation North American Fuel Survey©".[16]

4.2. Data Harmonization

4.2.1. Chemical composition data harmonization

Prior to statistical analysis, the chemical composition data were harmonized using a Python program that was originally developed within the CRC-sponsored project AVFL-29-2. In the current project, the Python program was adapted to include the GC-VUV datafile format, as well as to update the compound list and their properties. The compound list and species properties were updated using the PMI calculator v4.0.0.2.[17] Besides retention time (and retention index), parameters provided for each component included carbon number, hydrocarbon group type, vapor pressure, number of double bond equivalent (DBE), and the PMI factor. Compound group type is defined for each species as being in one of seven compound families. Besides the 5 usual classifications of n-paraffins, iso-paraffins, olefins, naphthenes, and aromatics (PIONA), group types are added for oxygenates (designated "X") and unknowns (designated "U"). We have also a list of synonyms to convert compound names to those in the Master List. For example, simple differences in terminology lead to different compound names, such as "isoprene" is often used instead of "2-methyl-1,3-butadiene," or "1,3-dimethylbenzene" is used instead of "m-xylene." The Python program recognizes such synonyms and coverts them to compound names used in the Master List. The program also recognizes more subtle discrepancies such as including a hyphen ("di-methyl") or not including a hyphen ("dimethyl"). The synonym assignment is done based on the synonym list as well as the compound retention index. All conversions performed by the program are recorded and thus can be checked or reverted if necessary.

In addition to the named (known) compounds, the Python program can estimate the PMI contribution of generic and unidentified compounds using their retention index.[15] While GC retention time is determined for each chromatographic peak, the retention index (RI) is commonly used for peak identification purposes. As GC columns age and other chromatographic parameters vary, retention times of eluting peaks tend to "wander," making peak identification more difficult. In contrast, RI values remain more constant, which increases the reliability of peak identification. The program uses the Kovats RI values – either as reported directly from the laboratories or, when RI values are not reported, as calculated by the program using the reported retention times. The formula for calculating Kovats RI values is shown below:

$$RI_i = 100\left(n + \frac{log(t_i) - \log(t_n)}{log(t_{n+1}) - \log(t_n)}\right) \quad (5)$$

where RI_i is the retention index of component i, n is the carbon number of n-paraffin, t_i is retention time of component i, t_n is the retention time of preceding n-paraffin, t_{n+1} is the retention time of the next n-paraffin.

Following the approach we adapted in the project AVFL-29-2, we used the RI of a compound to estimate its PMI factor, as well as its boiling point using regression formulas such as $\log_{10}$ (PMI) = a + b * RI. The regression is performed for each known compound within a PIONA class and then is used to estimate the parameters for generic compounds of that class. For unknown compounds, a regression is derived for all known compounds of all classes. Estimating the PMI contribution of generic and unidentified compounds helps to evaluate their contribution to the overall PMI and gain insight into any differences between different methods of PMI determination. For the D6730 summer dataset, this interpolation procedure resulted in PMI values that were 17% ±6% higher than PMI values calculated without interpolation, while for the winter dataset the increase was 20% ±11%. For the D8369, the increase was 20% ±10% for both seasons.

Fuel sample attribution to the fuel grade, PADD, etc., was done using the provided "Gage" sample numbers. Python dictionaries were constructed to link "Gage" numbers to the fuel identification, its grade, sample location, etc.

4.2.2. Distillation profile harmonization

The D86 and D7096 (SimDis) analysis results are reported in different formats and thus needed to be harmonized. D7096 reports temperatures required to evaporate certain volumes of fuel given in 1% increments, while D86 reports them for volumes separated in 5-10% increments. Thus, for comparing D86 and D7096 data, the latter were converted to the D86 intervals. For other analyses presented in this report, the D7096 data were used as is, i.e., with 1% increments. In addition to the differences in the format and resolution of the reported data, D86 data posed an additional challenge as they were provided without the "Gage" number designation. To assign "Gage" numbers, data on fuel grade, location, and manufacturer were used to find the matching "Gage" number. This resolved the vast majority of samples except for three winter samples that were assigned Gage numbers based on matching all fields except the brand field, which likely was misspelled, namely: 15704 (listed as "Murphy 7682", while only "Murphy" can be found in the Gage data), 15722 ("Hyvee" instead of "Hy-Veee" in the Gage description, where it is also misspelled), and 15734 ("Luk Oil" instead of "Luk").

The chemical composition data were also converted to distillation profiles. Each chemical composition datafile contains information on the content of individual compounds and their boiling points, either known or estimated using the approach described in Section 4.2.1. The per compound data were sorted in ascending order of the boiling point temperature. The cumulative sum of compounds content is then calculated, which provides a relationship between the temperature and the amount of fuel expected to be evaporated at that temperature, i.e., a simulated distillation profile. The data can then be binned into appropriate ranges to be comparable with the D86 or D7096 data.

4.3. Data Exploration and Statistical Analysis

All data processing, analysis, and plotting was done using Python (v.3.11.5) and its various numerical, statistical, and graphing libraries, such as the NumPy (v.1.24.3), Python for Data Analysis (Pandas, v.2.0.3), StatsModels (v.0.14.0), and MatPlotLib (v.3.7.2) libraries. PMI cumulative distribution data from the Honda 2010 study[4] were digitized from the plot provided in the paper using the Plot Digitizer (www.plotdigitizer.com).

5. Results

5.1. Summer dataset

5.1.1. Overview of the summer dataset

The number of fuel samples per grade and PADD region in the summer dataset are given in Table 1. The total number of samples is 169, with most being regular unleaded (100 samples), while premium unleaded and E15 had 44 and 25 samples respectively. PADD 1 was the most represented with 46 samples, while PADD 4 the least (17) samples. There were no E15 samples from PADD 5 and only one from PADD 4.

Table 1. The number of gasoline samples in the summer dataset per fuel grade and PADD region.

	PADD 1	PADD 2	PADD 3	PADD 4	PADD 5	*Total*
Regular Unleaded	28	20	20	12	20	*100*
Premium Unleaded	14	8	8	4	10	*44*
E15	4	11	9	1	0	*25*
Total	*46*	*39*	*37*	*17*	*30*	*169*

Table 2. Maximum retention indexes for the entire summer dataset as reported for D6730 and D8369 analyses.

	D6730-Gage (FID)		D8369 (VUV)	
	Named	All	Named	All
mean	1716.1	1747.8	1308.4	1311.1
std	162.5	172.3	78.8	77.4
min	1299.5	1302.7	1007.8	1007.8
25%	1563.0	1600.0	1270.6	1271.1
50%	1762.3	1787.2	1300.0	1300.3
75%	1874.7	1885.0	1383.8	1383.8
max	1963.9	2051.8	1433.8	1433.8

5.1.2. Chemical composition data intercomparison

The two methods used to determine gasoline composition in the surveys utilize gas chromatography for separating individual fuel components with the sequential detection by either FID (D6730) or VUV (D8369) detectors. It should be noted that the two methods of chromatographic separation also differ as they utilize different chromatographic columns, which is reflected in differences in the maximum RI reported by the two methods. Table 2 lists descriptive statistics for the maximum RI found in individual chromatographic files for the two methods for the summer dataset. The maximum retention index reported by the Gage version of D6730 for the summer fuels is 2051.8, while that for the D8369 is 1433.8. This corresponds to the elution region of C20 and C14 paraffins, respectively. The named (both identified and generic) compounds had generally lower RI than the maximum RI reported, indicating the

difficulty identifying heavier compounds that tend to have higher RIs. It should be noted that D8369 does not report any unidentified compounds, as the VUV detector can identify at least the compound class for a particular chromatographic peak based on its VUV spectrum, while FID cannot provide such information. The smaller RI range of the D8369 relative to D6730 indicates that the D8369 method would miss the heavier compounds that are detectable by the D6730. Since the heavier compounds tend to have a significant contribution to the PMI (which is indeed the case for this dataset, as will be shown later), such differences in the RI range may lead to biases in the measured PMI between the two methods. Another potential shortcoming of the VUV method is its lower sensitivity, which is approximately 10 times lower than that of the FID method.

A comparison of the total weight contributions of individual PIONAX groups reported by the two methods shows that, in general, they agree well with each other (Figure 5). Paraffins, isoparaffins, and aromatics compare very well between the two methods with the regression lines being close to 1:1 and R^2 of the fit being 0.98. The biases and variability between the reported total weight contributions for individual PIONAX groups can be better observed as frequency distribution of the ratios of the FID (D6730) and VUV(D8369) reported totals that are shown in Figure 6. For paraffins, isoparaffins, aromatics, and oxygenates, the standard deviation of the ratio is within approximately 4%. Olefins and naphthenes, however, show a much larger spread in the ratios. The average biases between the methods are small for paraffins and isoparaffins, being on average less than 3%, with the VUV method reporting higher values. For aromatics, the FID method was biased 5% higher relative to VUV. For oxygenates, the summer FID-reported values were biased 5% lower relative to the VUV method. The FID method reported mostly higher olefin and naphthene content, with deviations from the VUV method having a very large variability.

These observed differences in the total PIONAX contribution provide some insight into the relative performance of the two methods. The FID (D6730) method uses chromatographic retention time to identify compounds (in some cases, an independent, more specific mass spectrometric detection is used to verify the identity of a chromatographic peak), while the VUV (D8369) method also uses VUV absorption spectra to identify chromatographic peaks. While VUV spectroscopy is less specific than mass spectrometry, it does provide information on a likely compound class based on the features of the absorption spectrum. This is why the D8369 method does not report any unknowns, as all chromatographic peaks could be assigned at least some compound class, i.e., reported at least as a generic compound such as an aromatic-C10". The D6730 method, on the other hand does not have such an ability and does report a certain (though small, as will be discussed in the following text) content of unidentified compounds.

If the FID method were constrained to the VUV RI chromatographic range, the FID-reported PIONAX values would be lower relative to the VUV method due to the presence of the heavier compounds that fall outside of the VUV chromatographic range. The weight fraction of compounds in the FID data that fall outside the VUV range is quite small, 1.8% ±3.6%. Despite this potential bias, our analysis indicates that this is not the case for all compound classes. Paraffins, isoparaffins, and oxygenates reported by the FID method are lower than those reported by the VUV method. However, aromatics, olefins, and naphthenes reported by the FID method are higher. The underestimation of aromatics by the VUV method is quite remarkable as this

method is very sensitive and specific for this compound class. The smaller RI range of the VUV method cannot explain this bias, as within the overlapping RI range the VUV still reports approximately 4% less aromatics than the FID method.

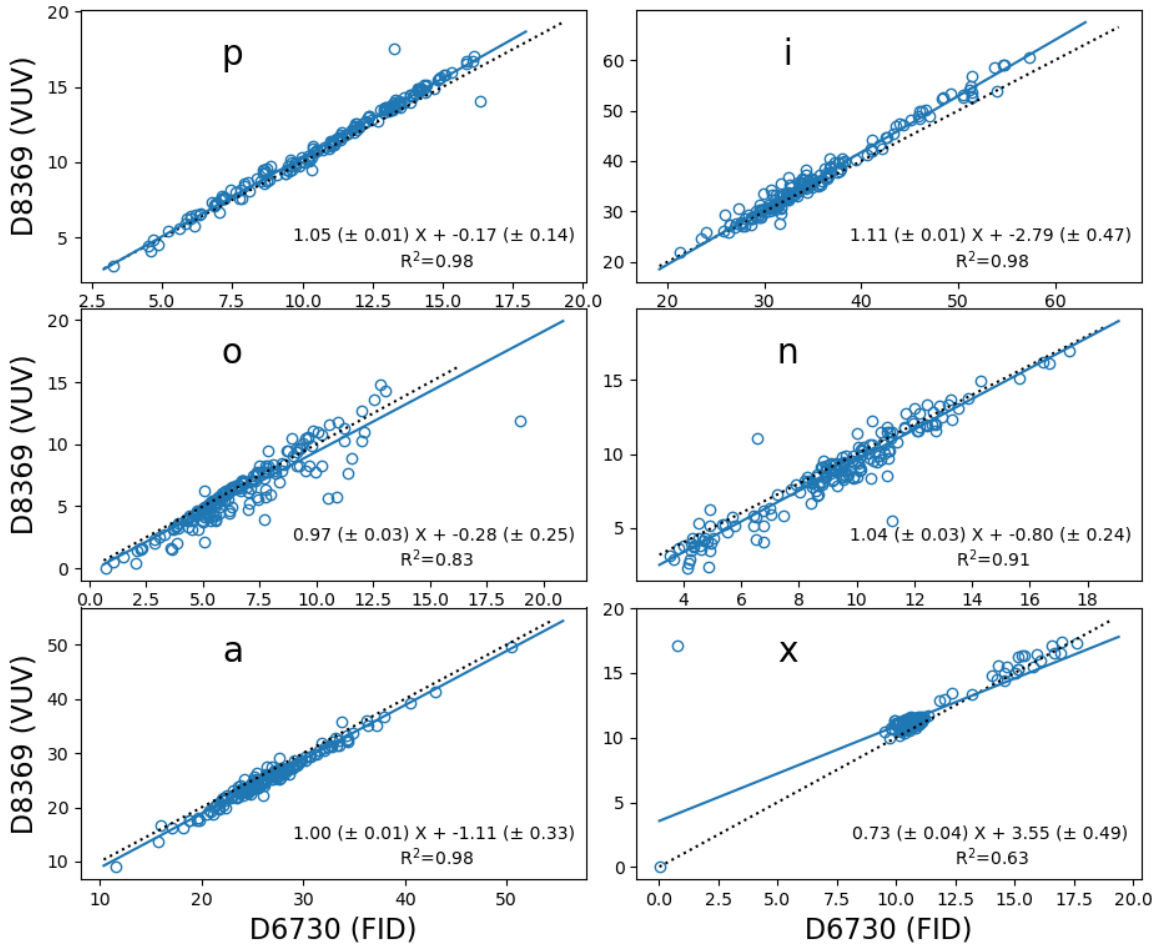

Figure 5. Comparison of total weight % of individual PIONAX groups reported by the D6730 and D8369 methods for the summer data set. The dotted line shows the 1:1 relationship, while the solid blue line shows the regression line with the regression equation and the corresponding R^2 of the fit given in the lower right corner of each subplot.

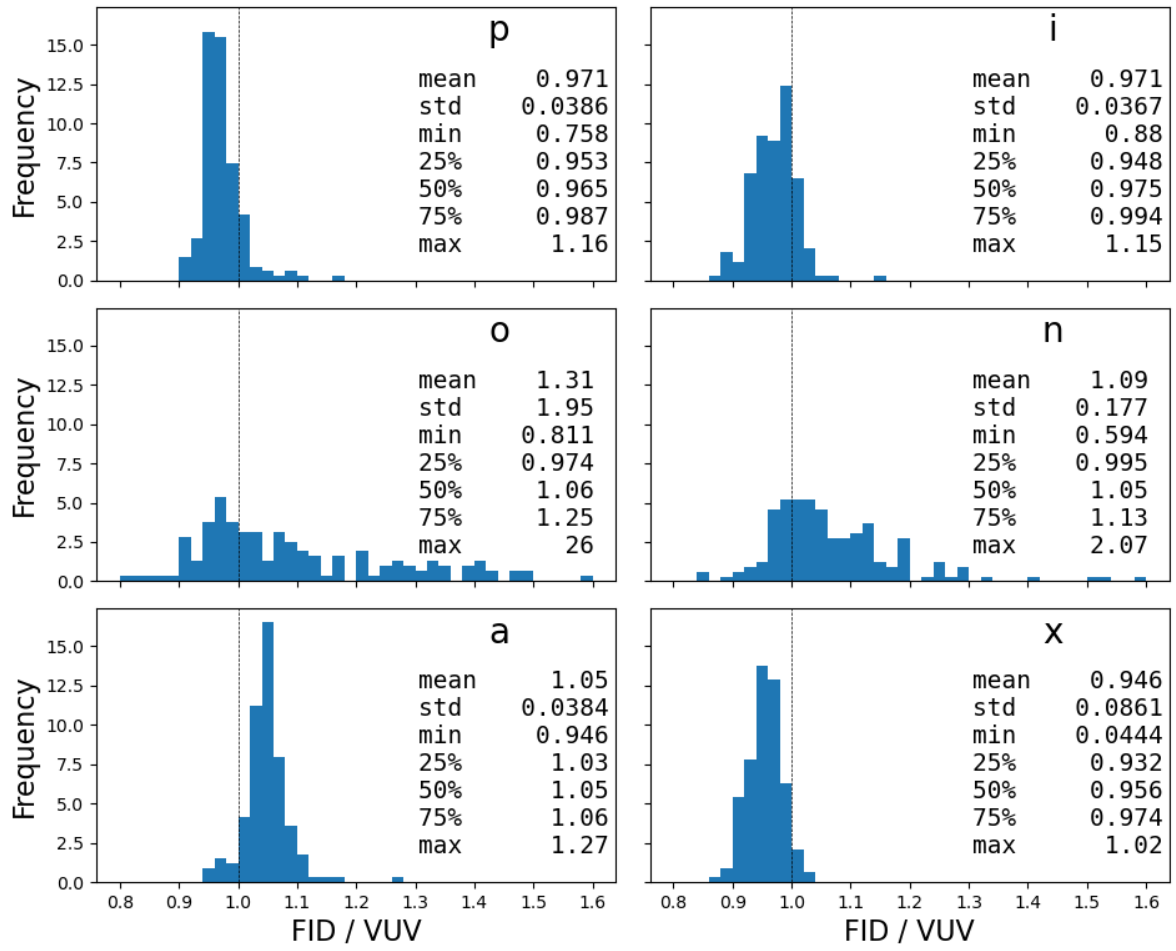

Figure 6. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the summer dataset. The descriptive statistics of the ratios are also provided.

A per-compound comparison, shown in Figure 7, provides further insight into the relative performance of the D6730 and D8369 methods. The figure was constructed by plotting weight percent of matching identified compounds for each fuel analysis. On average, the two methods agree reasonably well for compounds with individual weight contributions of approximately 1% or more, while for less abundant compounds the agreement is generally very poor, with the reported values often varying by one to two orders of magnitude. However, even for the higher concentration compounds, the agreement is often not perfect. For example, pentane content values correlate strongly between the two methods, but the FID method reports values that are 0.92 ± 0.02 of those reported by the VUV method. For heptane, the ratio of FID to VUV reported values becomes 0.98 ± 0.02, while for decane it is 1.52 ± 0.46, with lower concentrations having larger values. For benzene, the ratio is 0.93 ± 0.02. For naphthalene, it is 1.22 ± 0.33, while for all C10 aromatics (i.e., including substituted benzenes) the ratio is 2.1 ± 4.1.

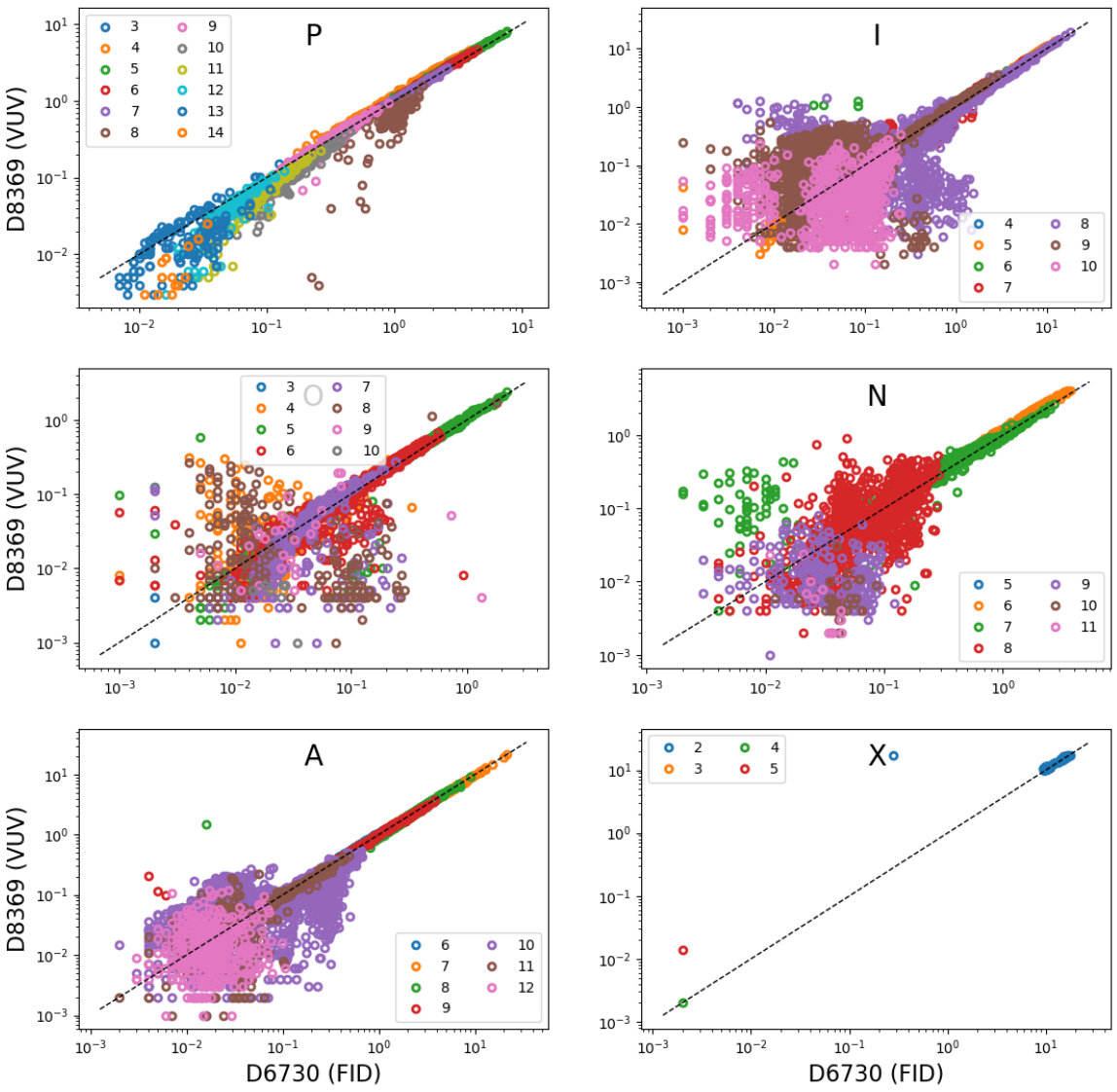

Figure 7. Per compound weight comparison within individual PIONAX groups reported by the D6730 and D8369 methods for the summer dataset. Compounds with the same carbon number are indicated by the same color (carbon numbers are given in the legends). The dotted line shows the 1:1 relationship.

As stated above, the lower aromatics readings reported by the VUV method are rather remarkable as it is supposed to be specific and sensitive for this class of compounds. This underestimation is especially pronounced for C10 aromatics (Figure 7). Coelutions could potentially bias some aromatics high, such as coelution of benzene and 1-methylcyclopentene or of toluene and 2,3,3-trimethylpentane. The VUV method uses peak deconvolution, which should, theoretically, separate any coeluting peaks. The FID method must rely on good chromatographic separation to avoid coelution artifacts. To check if such coelutions could have biased the FID measurements, we compared the concentrations of the above coeluting pairs reported by the VUV and FID methods, which is shown in Figure 8. No evidence can be observed for a coelution interference for the FID method. The FID reported 0.9% ±2% more toluene and 6.5%

±2.2% less benzene than the VUV method, while reporting more of both 2,3,3-trimethylpentane and 1-methylcyclopentane (109% ±177% and 20% ±26%). In fact, it appears that the VUV method could have suffered from coelution interference in benzene measurements. The FID method has a clearly superior chromatographic separation relative to the VUV method. The FID RI for benzene is 648.3 ±0.07 and 647.1 ±0.2 for 1-methylcyclopentane (i.e., 1.2 RI difference). These numbers for VUV are 646.56 ±0.18 and 646.46 ±0.92, respectively (only 0.1 RI difference). For toluene and 2,3,3-trimethylpentane, the FID RI is 750.5 ±0.3 and 749.2 ±0.2, respectively (1.3 RI difference). For the VUV method, these numbers are 749.41 ±0.74 and 749.79 ±0.9, only 0.38 RI difference. While the VUV method should be able to resolve coeluting peaks using deconvolution, it is not clear how accurate the deconvolution procedure is, especially at low concentrations, where the contribution of baseline noise becomes substantial.

Thus, the differences between the methods cannot be attributed to coelution artifacts. The differences at low concentrations are likely due to the poor sensitivity of the VUV method. Further, the fact that the abundant compounds such as pentane demonstrate consistent differences between the two methods indicates that there is a calibration bias between the two methods.

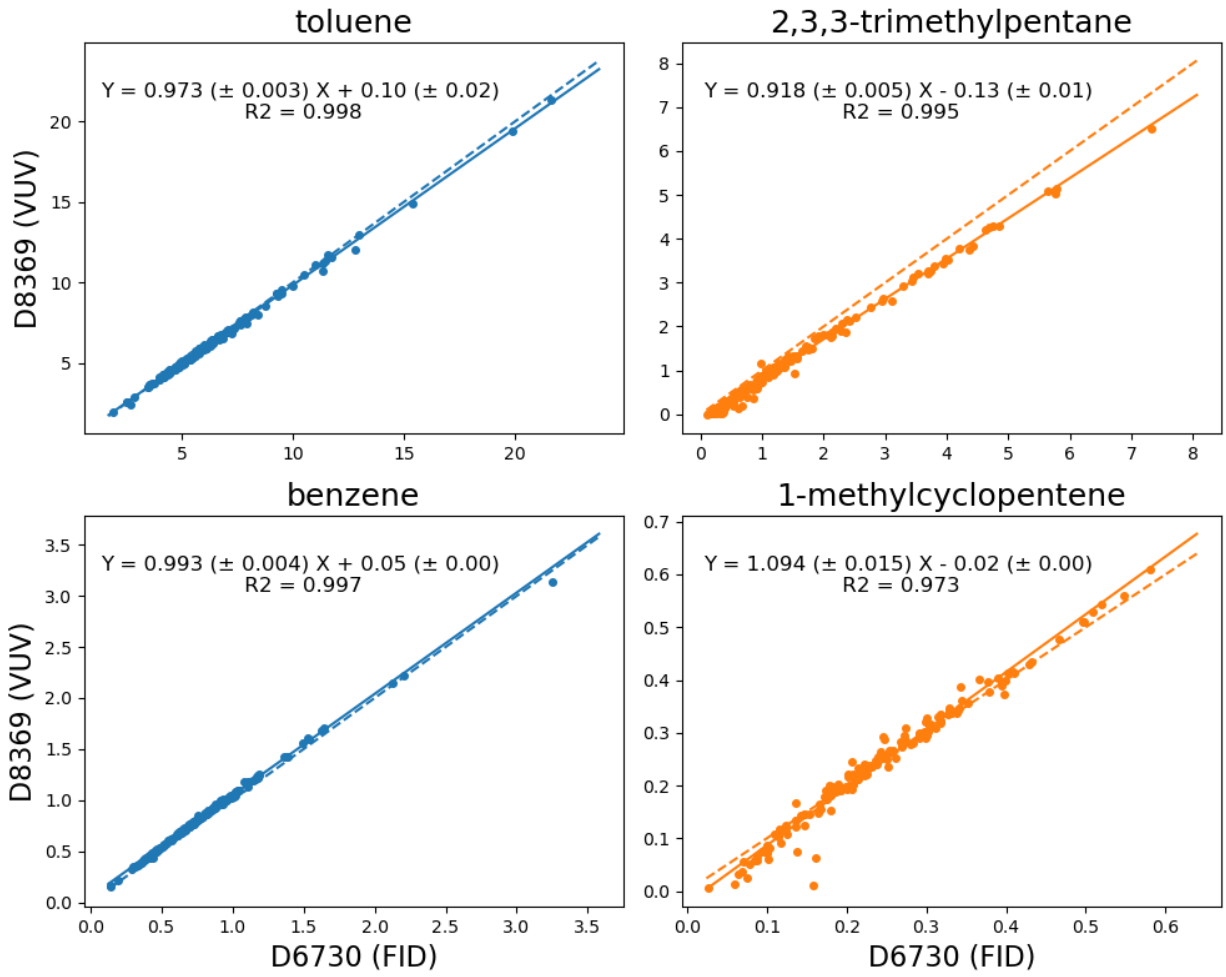

Figure 8. Comparison of concentrations of two potentially coeluting pairs (toluene and 2,3,3-trimethylpentane, and benzene and 1-methylcyclopentene) reported by the D6730 and D8369 methods for the summer dataset.

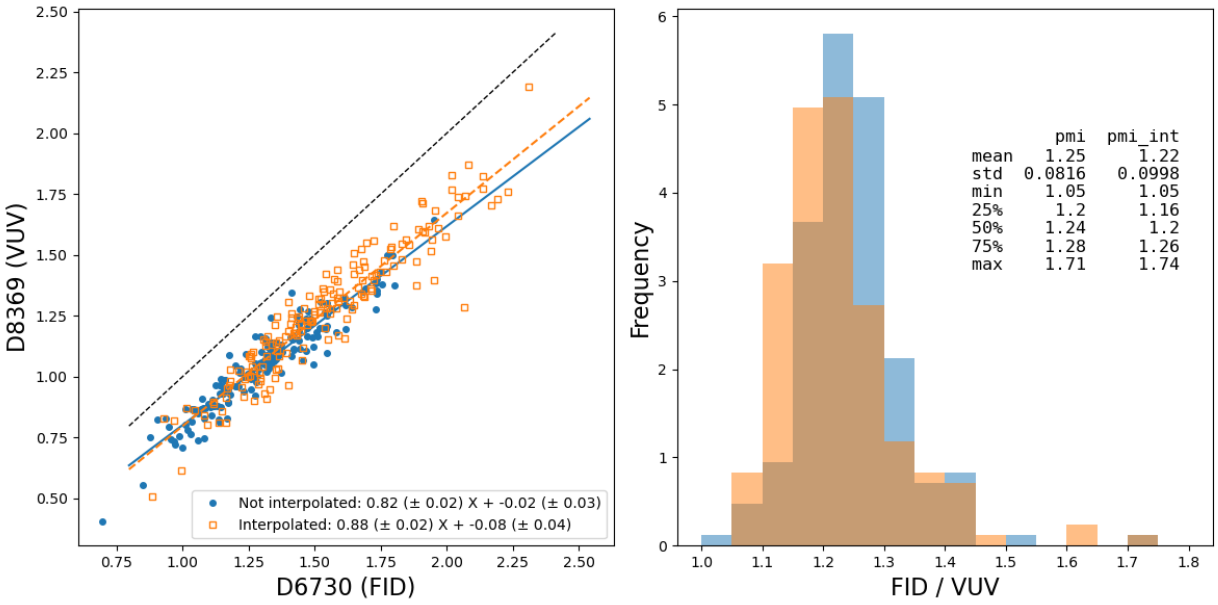

Figure 9. Left: Comparison of PMIs reported by D6730 (FID) and D8369 (VUV) methods for the summer dataset ("Not interpolated" – PMI data as reported, "Interpolated" – including interpolated PMI data). The broken black line represents a 1:1 line. Right: Histogram of FID/VUV ratios (the insert provides descriptive statistics of the ratios; "pmi" – PMI data as reported, "pmi_int" – including interpolated PMI data).

The reported total PMI values correlate well with each other, though with a consistent bias, with the D6730 reporting on average 25% higher PMIs than the D8369 (Figure 9). Including PMI values that were interpolated for generic and unknown compounds, improves the comparison somewhat, with the ratio becoming 1.23 on average. The underestimation of the total PMI by the D8369 method is larger than its underestimation of the weight contribution of the aromatics (1.05). As was discussed above, the underestimation of the weight contribution by the VUV method could be both due to biases in calibration between the two methods and the lower sensitivity of the VUV method.

5.1.3. Summer distillation profile intercomparison

Average distillation profiles measured with the D86 and D7096, as well as those derived from the D6730 and D8369 data, are shown in Figure 10. All 4 methods agree fairly well with each other between approximately 80ºC and 200ºC. The D6730 and D8369 agree well with each other at all temperatures, except for the highest temperatures (270ºC and 370ºC) due to the smaller chromatographic range of the D8369 method. The D86 and D7096 agree generally well with each other in the middle temperature range, with deviations increasing towards the lower and higher temperatures. In general, the D86 profile shows a later start of evaporation and an earlier end of evaporation relative to D7096. The D7096 indicates a much more volatile light end of the range than the other three methods, including the chromatography-based D6730 and D8369. It should be noted that ethanol content could influence the performance of individual methods. For example, unlike the chromatographic techniques, D86 profiles are sensitive to azeotrope effects, while the chromatography-based techniques are not. Within the chromatographic techniques,

there could be differences due to the column characteristics. For example, ethanol can potentially elute faster in the shorter, wide-bore column of the D7096 method, co-eluting with lighter gasoline compounds, which could explain the observed differences between the D7096 and composition-derived profiles. More details could be observed in a per-sample comparison that is shown in Figure 11.

Figure 11 shows violin and box plots of the per-sample differences in reported temperatures required to evaporate a certain amount of fuel. Fuel fractions were chosen to correspond to those reported by the D86 method. Violin plots show an estimate of the probability density function of the underlying data. Box plots are overlayed on top of the violin plots to indicate the main statistical parameters of the distributions, such as the range of data and outliers, inter-quartile range, and the median. The main features are similar to those discussed regarding the average distillation profiles. For example, the deviations are largest at each end of the spectrum. What is interesting to note is the presence of bi-modal distributions at some of the temperatures, which appears as two or more "bulges" on violin plots. The reason for these grouping is not clear as a cursory investigation of possible chemical composition differences that could cause such deviations did not reveal any obvious culprits.

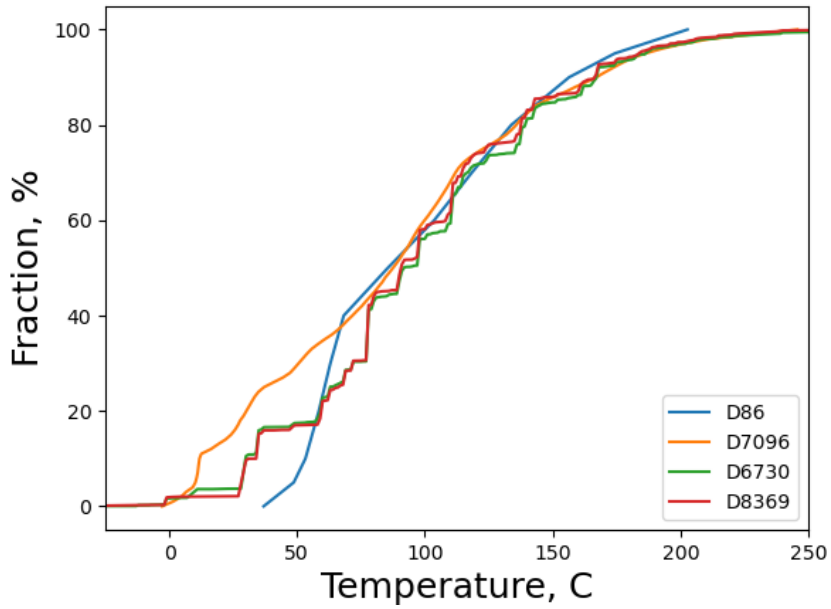

Figure 10. Comparison of the average distillation profiles measured with the D86 and D7096 methods and derived from the D6730 and D8369 data for the summer dataset.

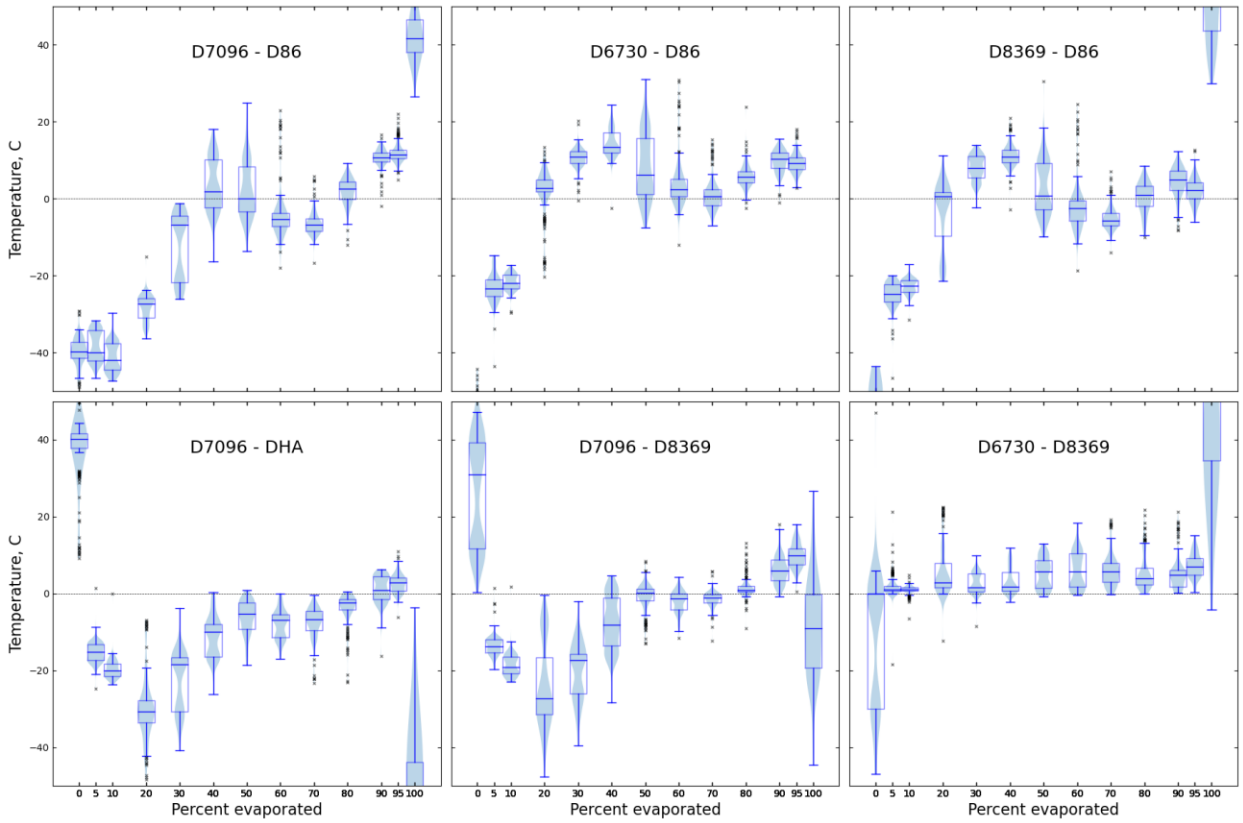

Figure 11. Violin plots and boxplots of per-sample differences in reported temperatures required to evaporate a certain fraction of fuel for the summer dataset.

5.1.4. Group contributions to weight and PMI

The average per group composition for the summer dataset is shown in Figure 12 and summarized in Table 3. The most prominent groups that contribute by weight are isoparaffins and aromatics, about 35% and 27% (based on the D6730 method), respectively. Paraffins, naphthenes, and oxygenates contribute approximately 10% each, while olefins contribute approximately 7% by weight. As discussed in Section 5.1.2, the VUV-reported isoparaffin values are biased higher relative to the FID method on average by about 3%. Approximately the same bias is observed in the season average (3.6%). Similarly, FID-reported aromatics are about 5% higher than those reported by the VUV method. The same trend holds for the other compound groups. The apparent differences in the average chemical composition between the two methods are thus likely to be a result of their consistent bias in either calibration or other method-specific differences.

Aromatic compounds dominate the overall PMI, contributing approximately 88% according to both the FID and VUV methods. Not including interpolated values for the generic and unknown compounds increases their contribution to 89%. Isoparaffins are a remote second PMI contributor (approximately 5% of the total PMI) with naphthenes being the third largest contributor (approximately 3%). Olefins and paraffins contribute approximately 1% each. The variability of the group weight and PMI contributions is shown in Figure 13 using the FID data.

Figure 14 shows the contribution of aromatic compounds to weight and PMI for the D6730 summer dataset, with the PMI values including those interpolated for the generic compounds. On average, C10 aromatics contribute most to the PMI, with the C8-12 groups all being very significant contributors to the PMI. C13 and C14 aromatics are relatively small contributors, though one fuel (Gage # 14798) had a very high C13 contribution comparable to the high values found for C8-C12 contributions. In any case, the distribution of PMI contributions over the aromatic carbon number shows that within the summer dataset the heaviest aromatics (C13+) contribute relatively little to the overall PMI.

Table 3. Average contributions of PIONAX groups to weight and PMI for summer dataset. "PMI interpolated" includes interpolated PMI values for generic and unknown compounds.

	Weight %		PMI		PMI interpolated	
	FID	VUV	FID	VUV	FID	VUV
group						
p	10.9	11.2	0.02	0.02	0.02	0.02
i	35.3	36.5	0.06	0.05	0.08	0.06
o	6.7	6.2	0.02	0.01	0.02	0.01
n	9.1	8.7	0.04	0.03	0.05	0.04
a	26.7	25.5	1.17	0.94	1.35	1.13
x	11.1	11.7	0.01	0.01	0.01	0.01
u	0.2	0.0	0.00	0.00	0.01	0.00

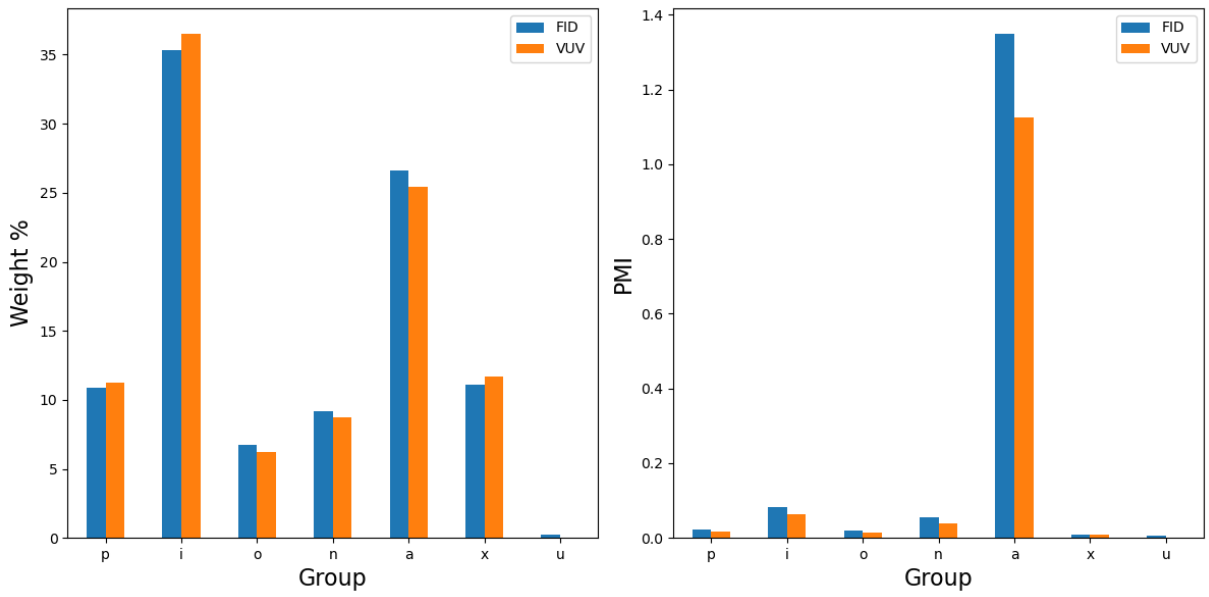

Figure 12. Average contributions of the PIONAX groups to fuel weight composition and PMI for the summer dataset.

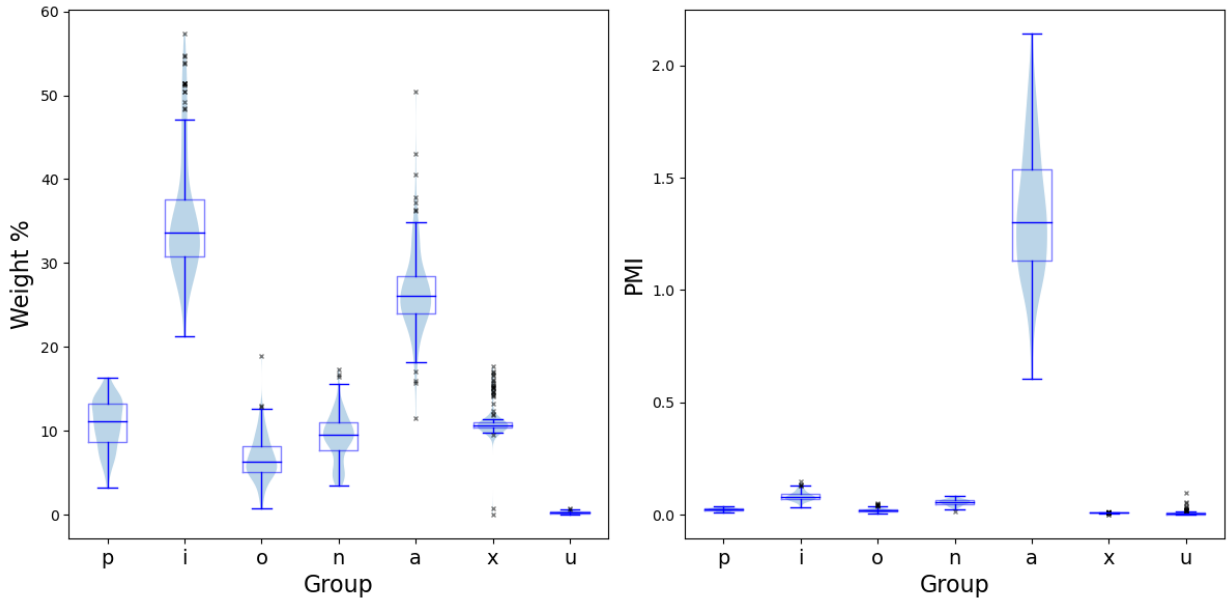

Figure 13. Compound group contributions to weight and PMI for the summer dataset based on the FID (D6730) data.

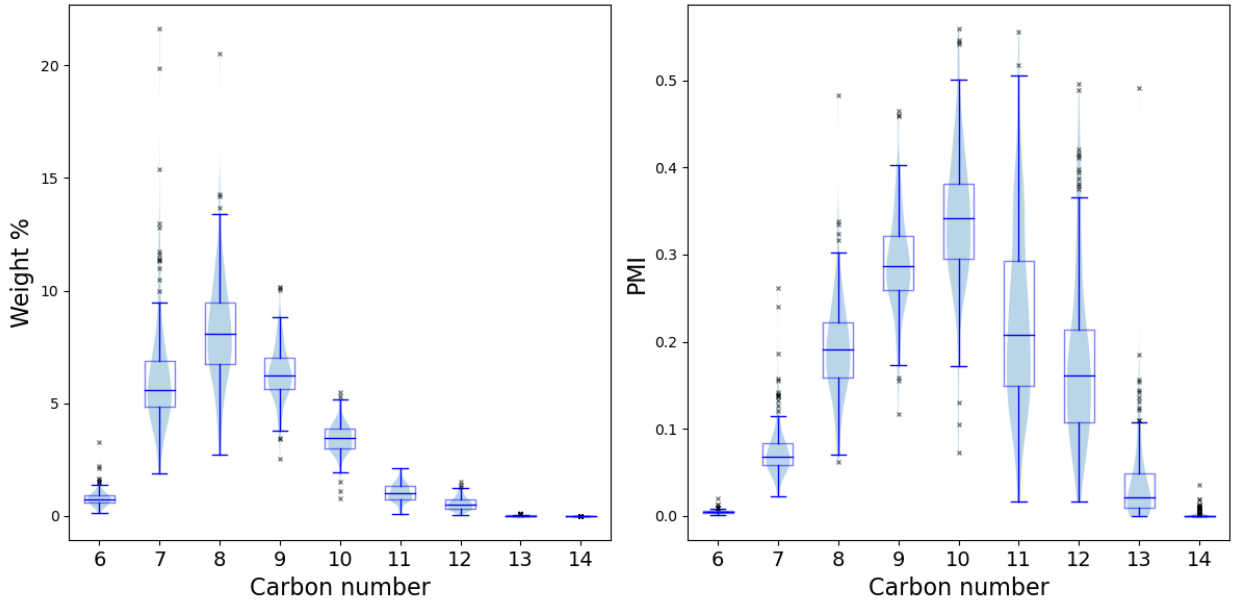

Figure 14. Contribution of aromatic compounds to weight and PMI for the D6730 summer dataset. PMI values were calculated including interpolated values for generic compounds.

5.1.5. Comparison of summer gasolines across grades and PADD regions

A breakdown per group and carbon for the three gasoline grades is shown in Figure 15, which shows the average contributions of compounds with different carbon number for individual

compound groups over the entire summer dataset. The most apparent differences in chemical composition between the grades, aside from the higher ethanol content of the E15 grade, are the higher C8 isoparaffin and toluene (C7) contents of the premium unleaded grade relative to the other two grades. This is consistent with isooctane and toluene being used to obtain the higher (R+M)/2 for premium grade gasoline.

However, there are statistically significant differences in other compound groups among all three grades. Figure 16 shows a comparison of the three grades per group and carbon number. The figure indicates the statistical significance of the differences based on the Wilcoxon signed rank test. This is a non-parametric test that does not depend on the normal distribution assumption and can handle the unequal distribution of samples between the three fuel grades. The statistical significance is indicated in the following way: three stars denote p values below or equal to 0.001 (very high significance), two stars: $p < 0.01$, one star: $p < 0.05$, and "ns" for p values above 0.05 (not statistically significant at 5% confidence level). The premium unleaded gasoline differs from the other two grades in almost all major components (excluding ethanol), probably because the high content of C8 isoparaffins brings all the other compound content down. The regular and E15 gasolines are very similar in most, except ethanol, compound classes. In addition to ethanol, however, there are statistically significant differences between the two grades in C7 and C8 paraffins (heptane and octane), C13 and C14 isoparaffins, C7 olefins, C6 naphthenes, and C7, C11, C12 aromatics.

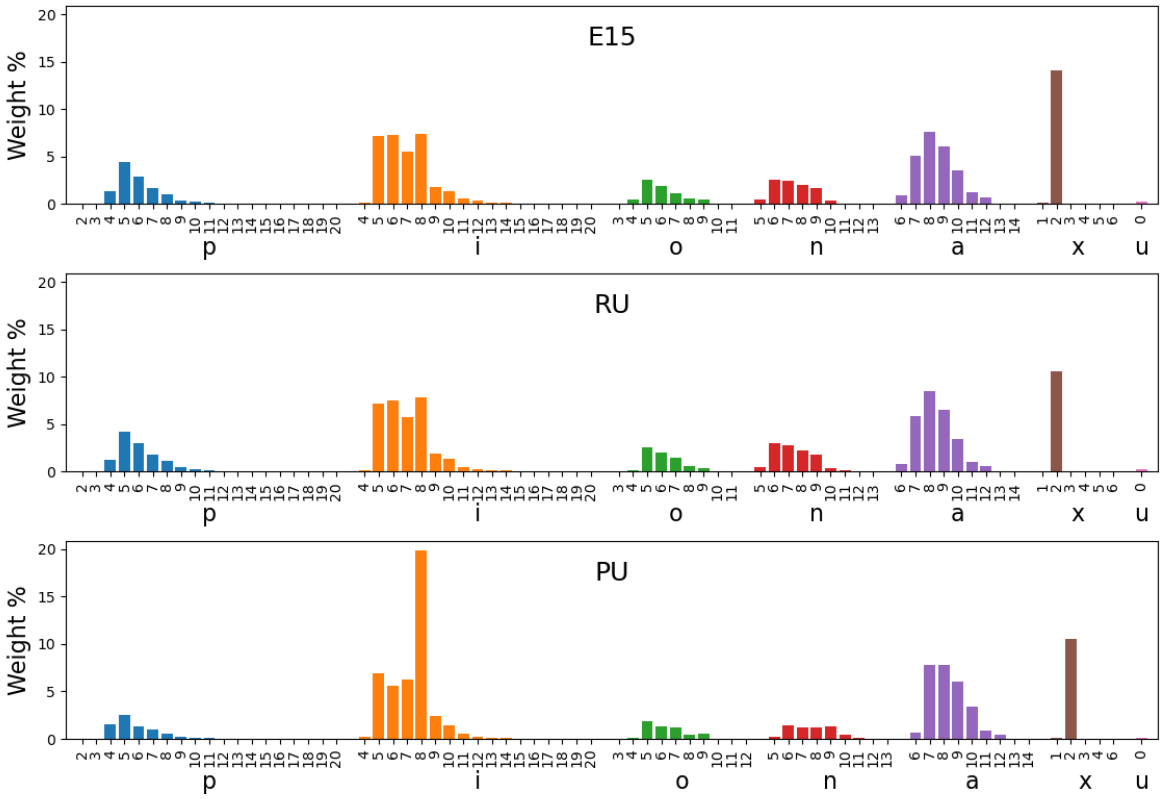

Figure 15. Average composition of the three gasoline grades per PIONAX group and carbon number based on D6730 summer data (RU: regular unleaded, PU: premium unleaded).

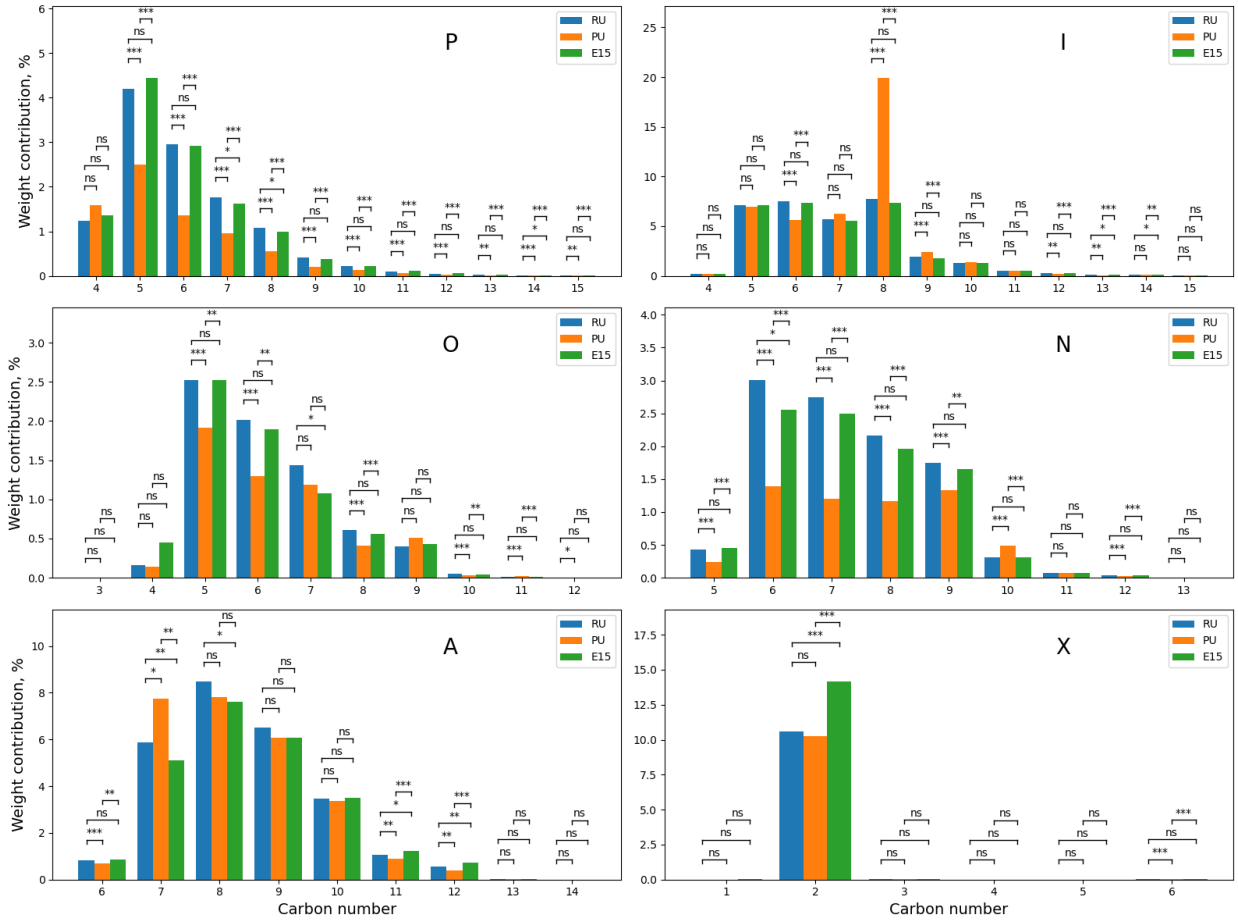

Figure 16. Statistical comparison of chemical composition of the three fuel grades using D6730 summer data (RU: regular unleaded, PU: premium unleaded). Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

Statistically significant differences between different PADD regions were also observed. Similar to the analysis presented above, fuels of the same grade were compared across different PADDs using the Wilcoxon rank test. Figure 17 shows such a comparison for premium unleaded gasolines. For most of compound subclasses there are no statistically significant differences between any of the PADD regions. However, there are several differences that are significant. For example, C4 paraffins are higher in PADD 4 than in PADD 5, pentane (C5) is higher in PADD 1 than in PADD 5. C6-C9 naphthenes are higher in PADD 5, while C7 aromatics tend to be lower in PADD 4 and 5 relative to the other regions.

Among regular unleaded fuels, there are statistically significant differences between different PADD regions mainly in C4-C6 paraffins, C4-C8 isoparaffins and olefins, C6-C7 naphthenes, C6, C11 and heavier aromatics (Figure 18). A comparison of E15 fuels was only possible for PADD 1-3, as PADD 4 had only one sample of that grade and PADD 5 had none (Table 1). This comparison is shown in Figure 19. Pentane content in PADD 3 is lower than in the other two regions, while heptane and octane are higher. C4 and C5 isoparaffins are lower in PADD 3 relative to PADD 1 and 2, while C8 and some heavier isoparaffins are higher in PADD 3 than in

PADD 1. There are also significant differences among some olefins, naphthenes, and C11-12 aromatics. The data also indicate that the ethanol content of E15 fuels in PADD 3 is statistically significantly lower than in PADD 2. It should be noted that the number of samples per grade and PADD is rather limited (see Table 1), which reduces the statistical power for small deviations. With a larger number of samples per each PADD, other differences could potentially become statistically significant.

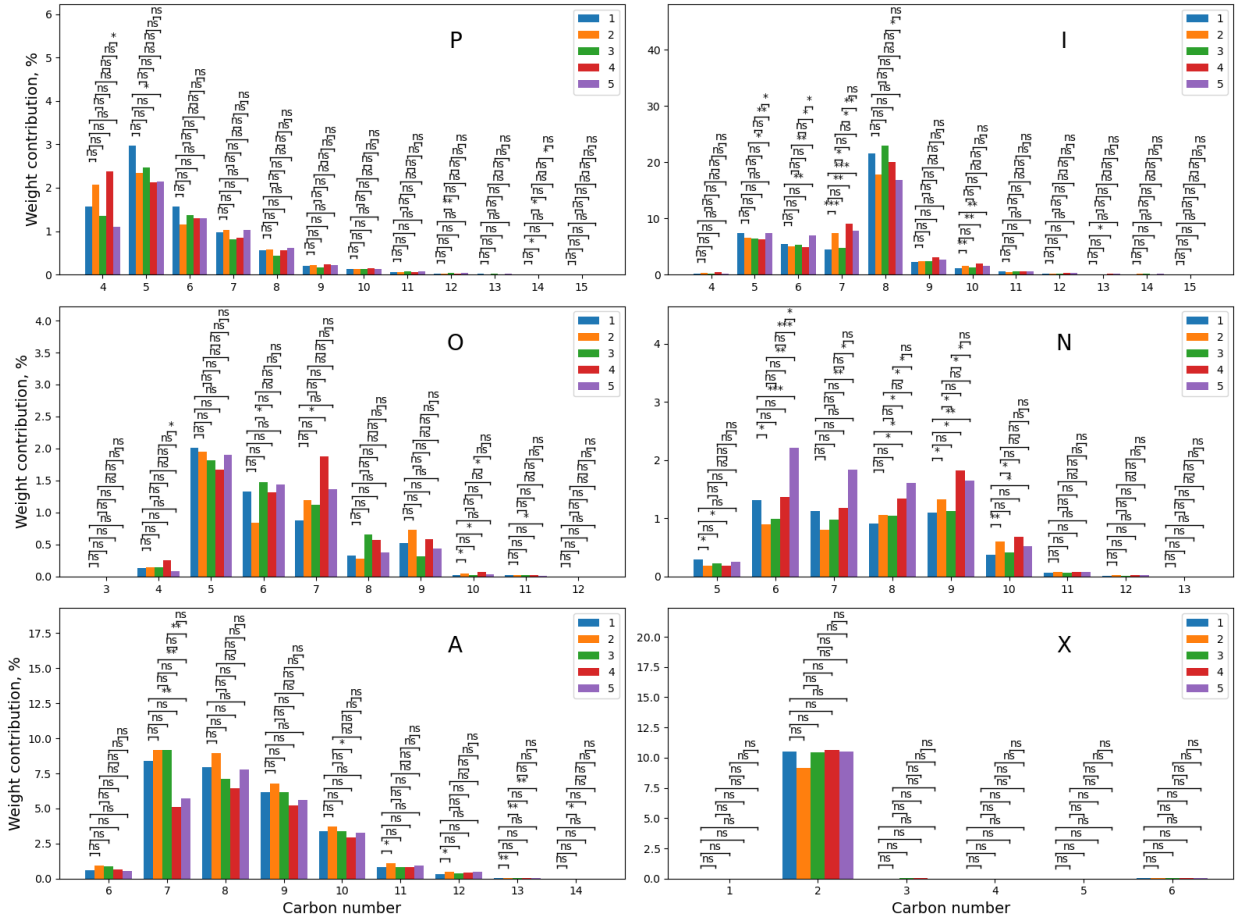

Figure 17. Statistical comparison of chemical composition of summer premium unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

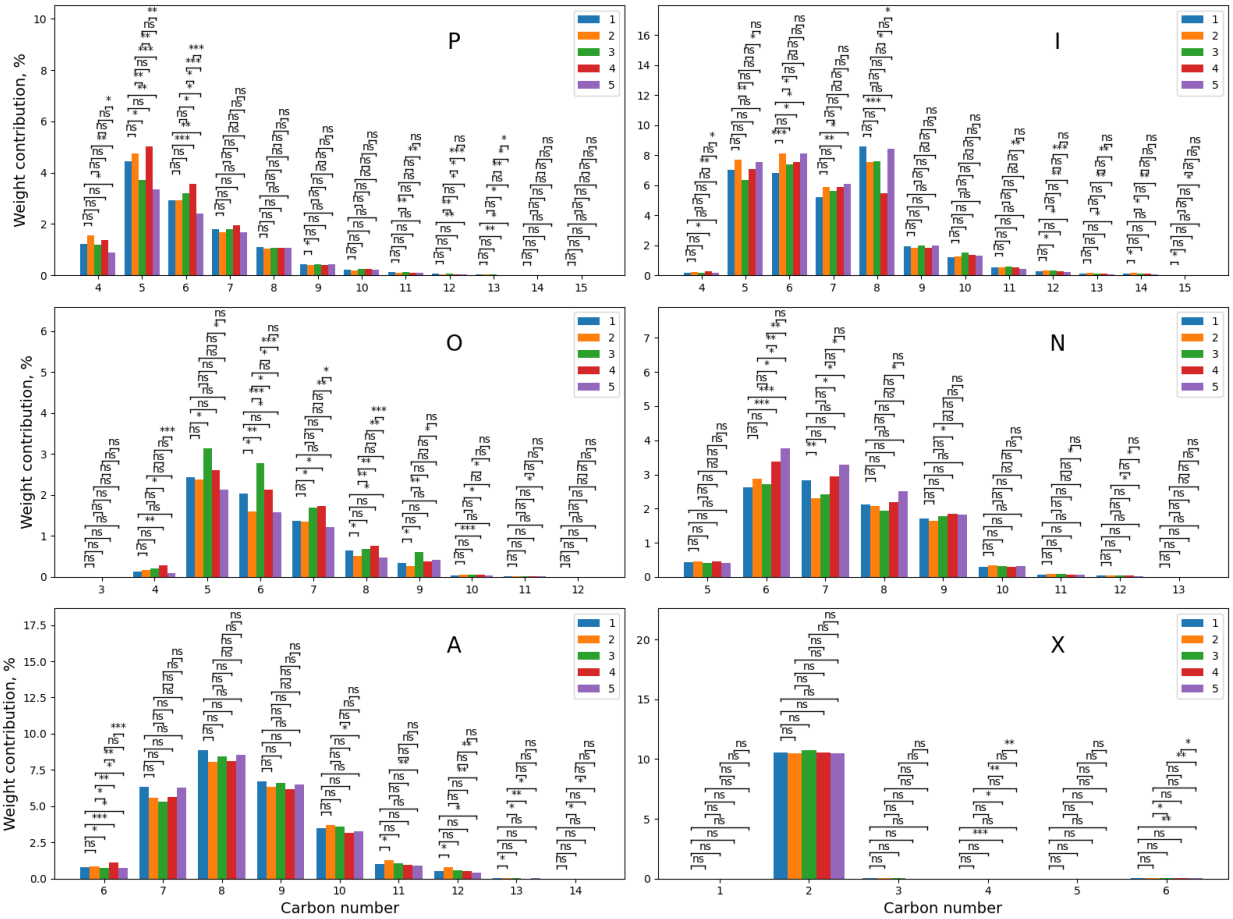

Figure 18. Statistical comparison of chemical composition of summer regular unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

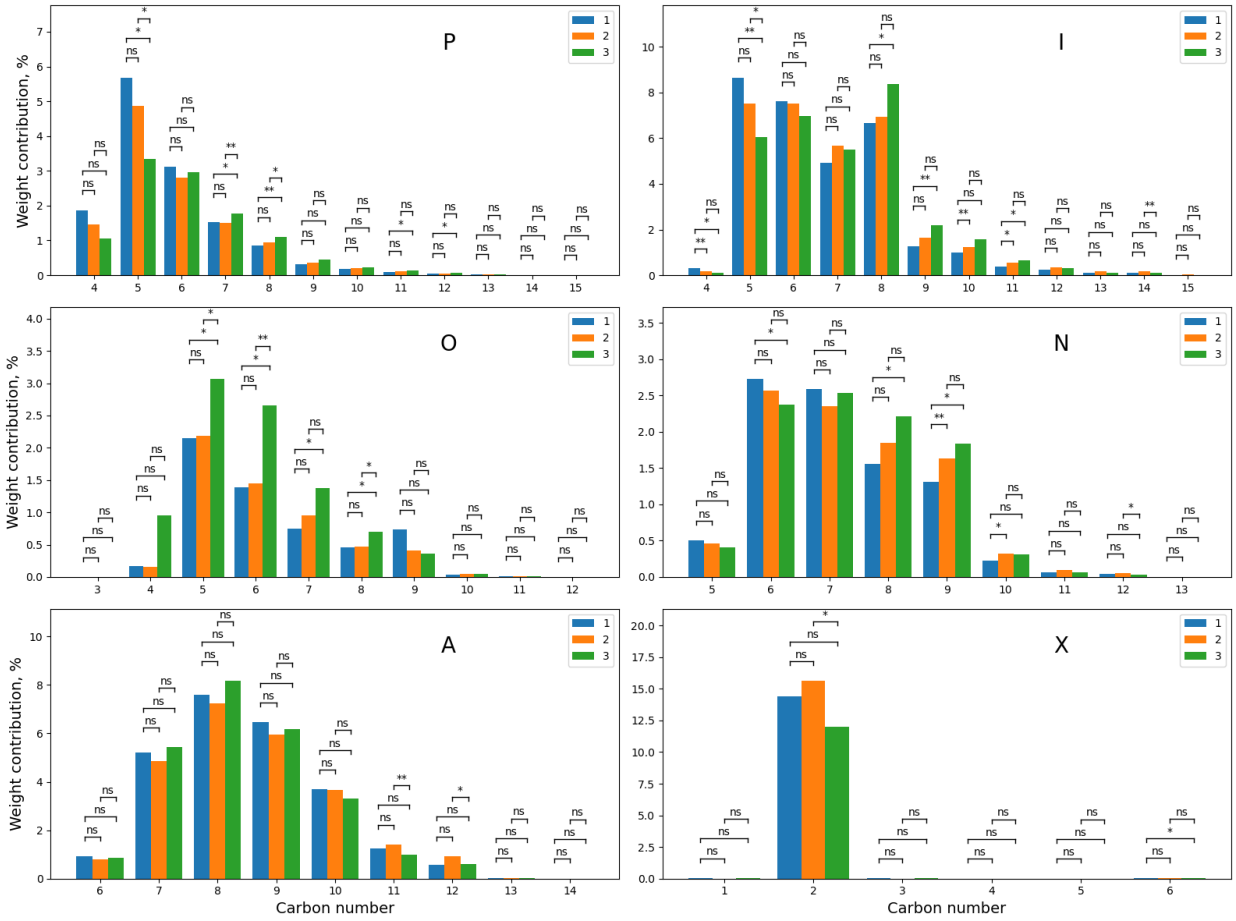

Figure 19. Statistical comparison of chemical composition of E15 gasoline in PADD regions 1,2,3 broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

5.1.6. Comparison of PMI across grades and PADD for summer dataset

A comparison of PMI values between the fuel grades and the five PADD regions for the entire summer dataset is shown in Figure 20. The figure is based on the D6730 analysis and includes interpolated PMI values for generic and unknown compounds. The Wilcoxon test was applied to assess the statistical significance. Among the three fuel types, the premium unleaded fuels are statistically significantly lower than the other two grades, while the differences between the regular and E15 fuels are not statistically significant. Fuels in PADD 2 have statistically significantly higher PMI values than fuels in the other regions, while the differences between the other regions are not statistically significant.

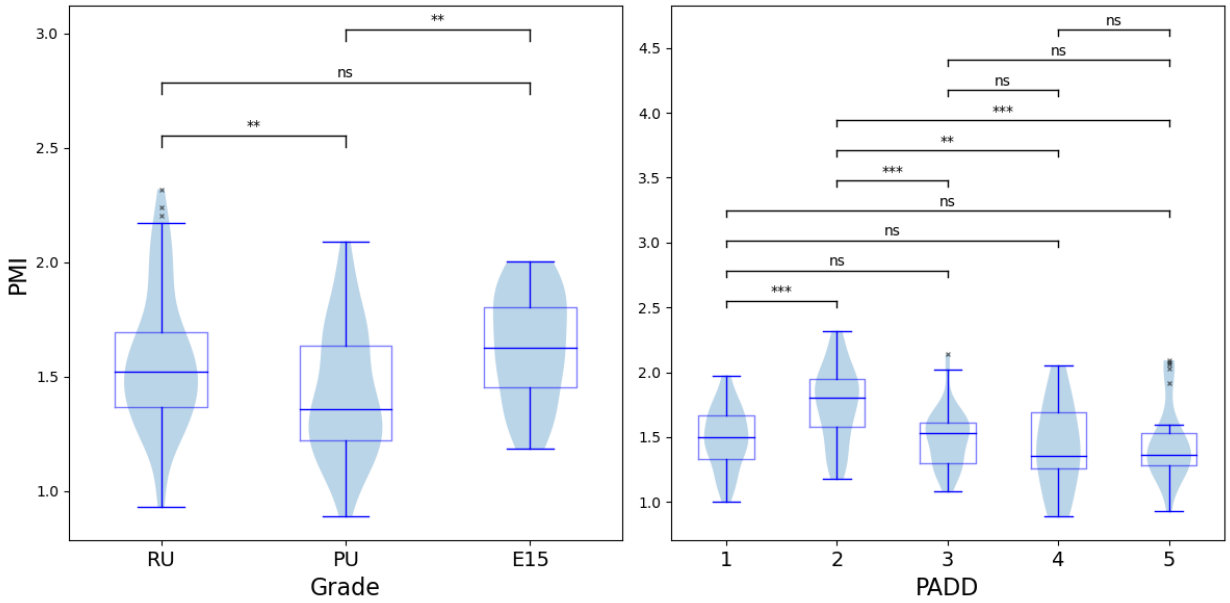

Figure 20. Statistical comparison of PMI of different grades of fuel and fuels in different PADD regions for the entire summer dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

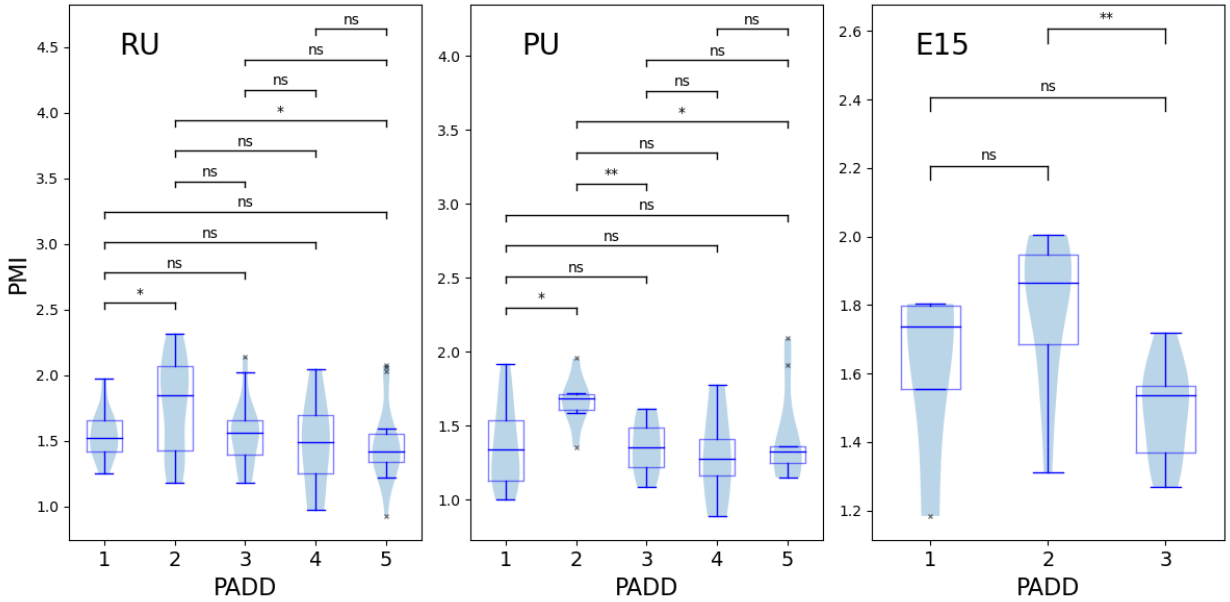

Figure 21. Statistical comparison of PMI between different PADD regions for the three fuel grades for the summer dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

Figure 21 shows a comparison of fuel PMI in different PADD regions for each of the fuel grades. Due to the absence of E15 fuel samples in PADD 5 and only one sample in PADD 4, it was not possible to include these regions in the comparison of E15 fuels. This breakdown shows that

PMI of regular unleaded fuels in PADD 2 are statistically significantly higher than those in PADD 1 and 5, while the differences with other and among other regions are not statistically significant. PMI of the premium unleaded fuels in PADD2 is statistically significantly higher than those in PADD 1, 3, and 5. E15 fuels in PADD 2 have statistically significantly higher PMI than those in PADD 3. No other statistically significant differences are observed.

5.1.7. Contributions of generic compounds to weight and PMI in summer fuels

The contribution of generic compounds, i.e., those for which only the group was identified, to the total weight contribution by PIONAX groups for the D6730 (FID) summer dataset is shown in Figure 22. The contribution was calculated as the ratio of the weight percent of generics to that of the total of each group. Paraffins and oxygenates are not shown as there are no generics reported in these groups. Likewise, all unidentified compounds are by definition unidentified. Overall, unidentified compounds contributed 0.23% ± 0.13% by weight, which demonstrates the high identification power of this DHA method. Aromatic weight is the most identified group, with only about 1% of all aromatics weight being generic. The naphthenes, on the other hand, had the highest mean of 12.8% of generic weight, with 3 fuel samples reaching above 25% generic content of naphthenes.

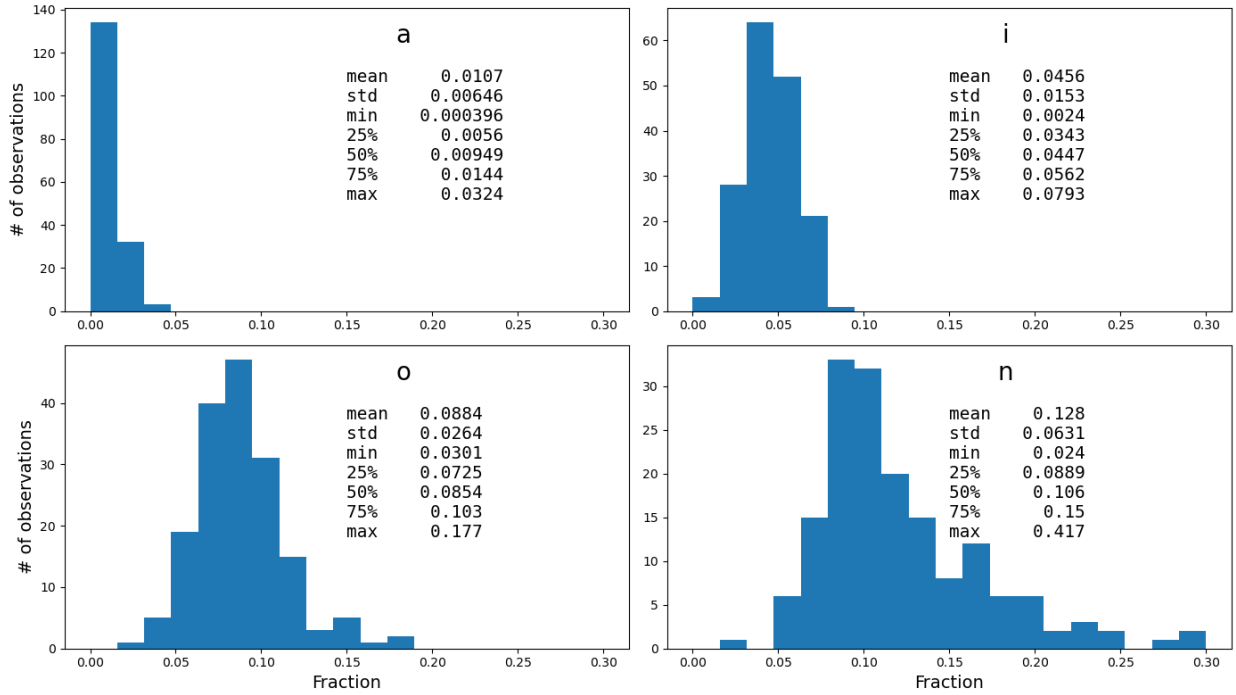

Figure 22. The weight fraction of generic compounds within aromatic ('a'), isoparaffin ('i'), olefin ('o'), and naphthene ('n') groups for the summer D6730 dataset (all grades and PADD).

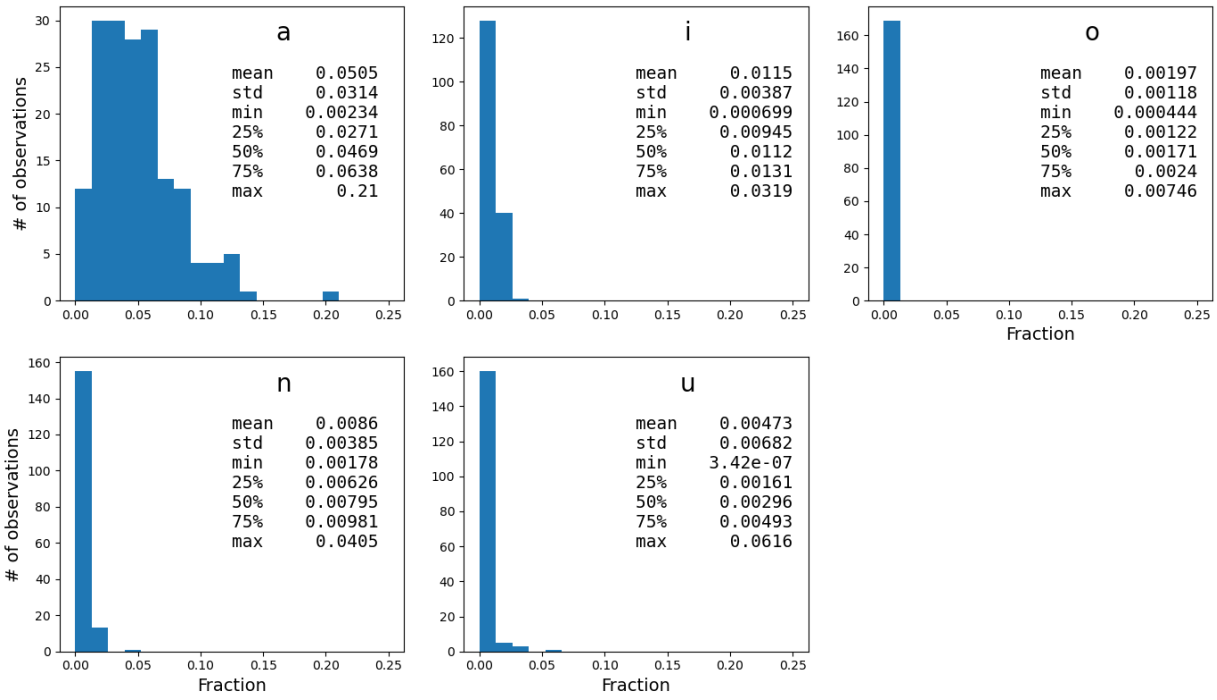

Figure 23. PMI contribution of generic and unidentified compounds to the total fuel PMI based on the D6730 summer dataset. 'a': aromatics, 'i': isoparaffins, 'o': olefins, 'n': naphthenes, 'u': unidentified.

The contribution of generics to the total fuel PMI (as opposed to the individual group weight contribution shown in Figure 22) is shown in Figure 23. This contribution is most pronounced for the aromatics compounds, where generics contribute about 5% ± 3% of the total fuel PMI, with one fuel (Gage # 14798) having a 20% contribution of generic aromatics to the overall PMI of that fuel. This fuel had a very high C13 contribution relative to other fuels (Figure 14), which suggests that C13 generic compounds could be the culprit of the high generic contribution for this fuel. It is also possible that there was a class misattribution by the chromatography software. Generic compounds in other classes contributed less than 1% of the total fuel PMI. Unidentified compounds contributed about 0.5% ± 0.7% of the total fuel PMI. This again testifies to the resolving and identification power of the D6730 method. The higher contribution of the aromatic generic compounds is likely due to their heavy contribution to the overall PMI, especially at the heavy molecular weight end. However, it should be noted that the contribution of generic aromatics to the overall PMI is rather small (5%) compared to other uncertainties involved in relating fuel properties (such as those expressed by PMI) to PM emissions by internal combustion engines. It thus appears that at the current state of DHA technology the chromatographic resolution is not the limiting factor in estimating PM emissions for a given fuel.

5.1.8. Summer distillation profiles

The average distillation profiles measured with D86 and D7096 as well as those derived from the D6730 and D8369 measurements are shown in Figure 10. To provide a measure of the variability between individual distillation profiles, Figure 24 shows violin and box plots for the summer dataset. As was discussed in Section 5.1.3, the biggest differences are observed for the lowest

and highest temperatures, where the FID method (D6730) has much larger tails due to its larger chromatographic range. Overall, the variability among individual distillation profiles is fairly small within each of the measurement methods. The D86 profiles have the greatest variability in temperatures required to evaporate 50%, while D7096 profiles are mostly variable in the 30%-50% range. The reason for such variability is not clear and their differences between the methods is not clear.

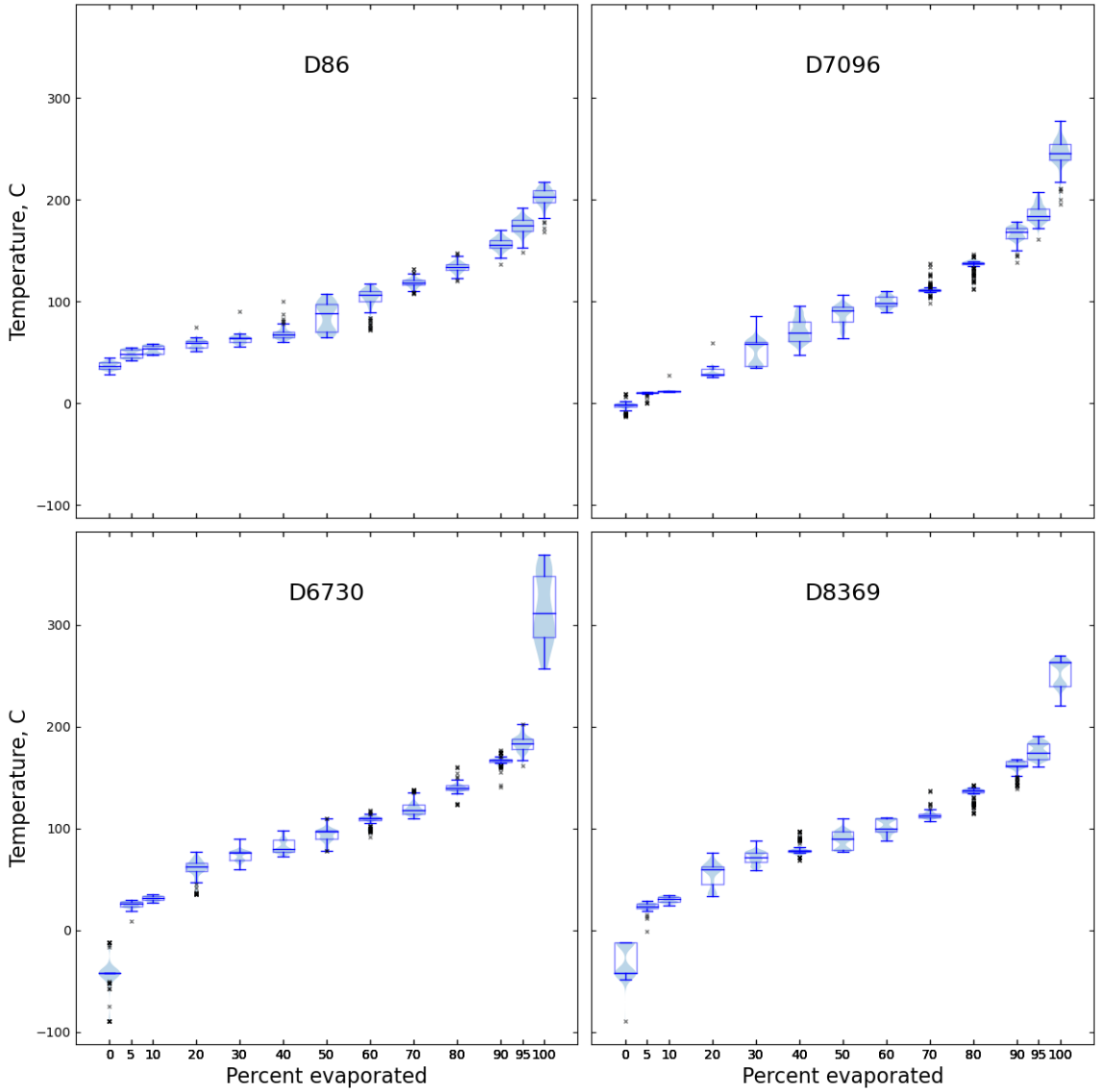

Figure 24. Violin and boxplots of distillation profiles (temperatures required to evaporate a certain fraction of fuel) for summer dataset (alld grades and PADD).

5.1.9. Hypothetical effect of distillation cuts on summer fuel PMI

Discarding the highest boiling fraction of a fuel (the so-called distillation cut) was proposed as a possible approach in reducing the fuel PMI and thus PM emissions. Indeed, there is a correlation observed between the fuel PMI and temperatures at which certain fraction of fuel is evaporated,

such as 95%, 97%, and 99% (Figure 25). The relationship between the temperature estimated for a certain distillation point and fuel PMI was determined in the following way. First, the temperature corresponding to a certain evaporation fraction was estimated by linearly interpolating, if necessary, the distillation profile, i.e., the relationship between the boiling temperature and the cumulative weight percent of fuel components. The method for obtaining distillation profiles from DHA data is described in Section 4.2.2. The temperatures corresponding to a certain cut were then plotted as a function of the fuel PMI. Since D86 and D7096 data do not provide fuel PMI, PMI values from the D6730 method were used for plotting. For D8369, PMI values obtained with D8369 data were used.

The observed positive correlation between the fuel PMI and "distillation cut" temperatures is not surprising given that the heaviest compounds have a strong contribution to the overall PMI due to their high boiling points and the high abundance of unsaturated bonds. The correlations, however, are not very strong, with R^2 values ranging between 0.41 and 0.69 (Figure 25). For all methods, except D86, there are horizontal groupings of points, which is especially pronounced for the VUV method (D8369). Such groupings indicate that there are several fuels with different PMIs that have very similar "distillation cut" temperatures. The reason for this is likely due to the uneven distribution of material and the corresponding PMI quantities over the boiling point space. It should be kept in mind that all methods, except D86, are chromatography-based. In the chromatographic methods, individual compounds are eluted in discrete quantities separately from each other (which is less pronounced in the D7096 method due to its relatively poor chromatographic resolution). This results in an uneven distribution of material over the boiling point space and a stepwise shape of the derived distillation profile that is apparent even in the season-averaged profiles shown in Figure 10. The D8369 method has generally fewer compounds identified, especially in the heavy end, which exacerbates this phenomenon. The D86 method, on the other hand, is a distillation method in which fuel components evolve with temperature gradually according to their chemical activity in the remaining mixture at each temperature, not at their sharply defined boiling points. The D86 distillation profiles are thus smoother than those of the other methods, though the much lower temperature resolution of D86 could also be a contributing factor. It should also be stressed, especially with respect to the chromatography-based methods, that real-life distillation processes are much less chemically resolved. Thus, the estimates of distillation cut effects on fuel PMI presented here are likely to differ from real-life distillation cuts.

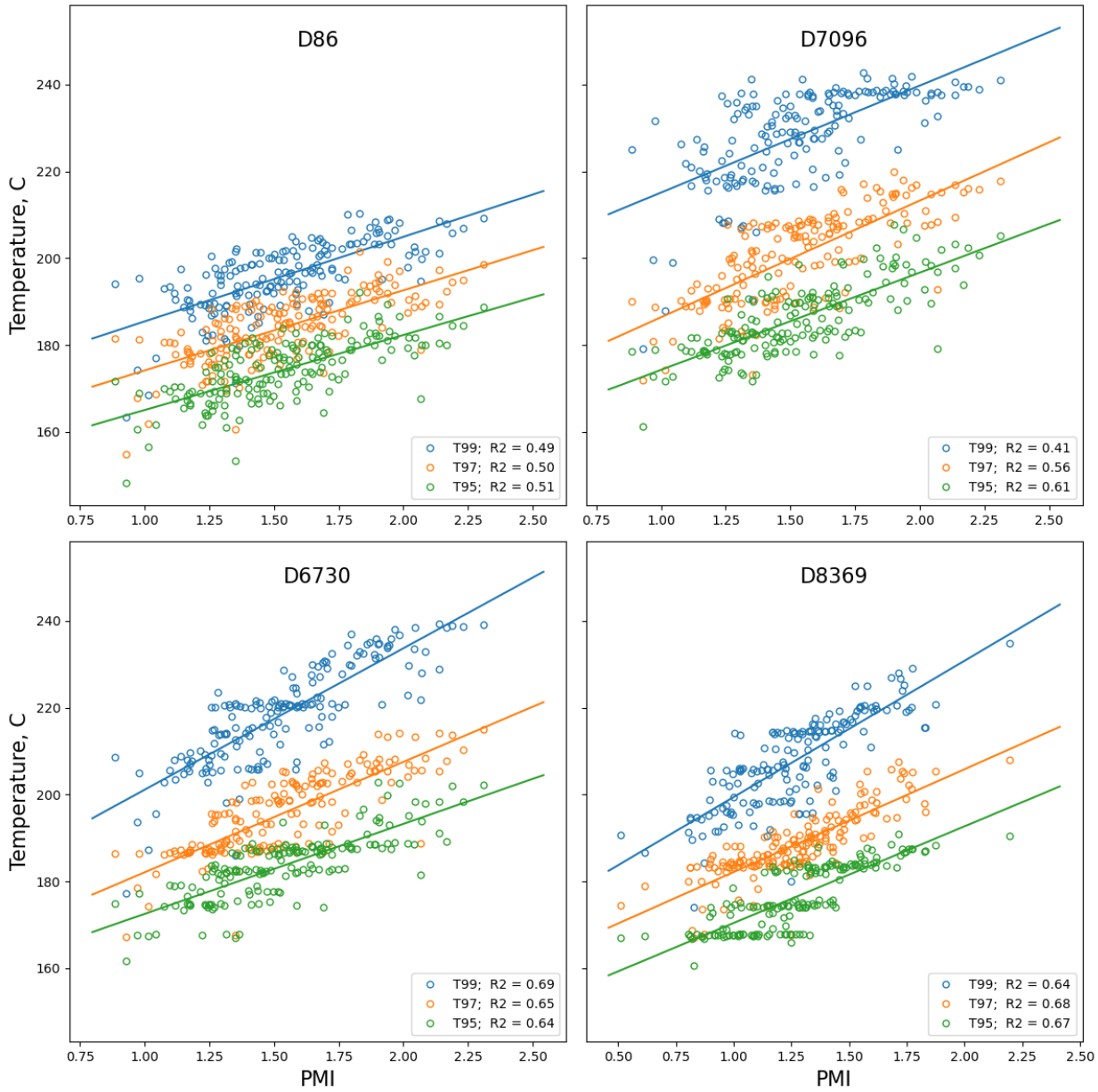

Figure 25. Temperature at which 95% (T95), 97% (T97), and 99% (T99) of fuel is evaporated vs. fuel PMI for the summer dataset for the four measurement methods. The D8369 subplot uses PMI values derived using the D8369 method; the other plots use PMI values derived using the D6730 method.

An estimate of how a distillation cut could affect a fuel's PMI was made in this project using the D6730 data by performing the following steps. The per-compound composition data (weight % and individual PMI contributions) were ordered in ascending order of the compounds' boiling points. Then cumulative values for the PMI were calculated as a function of the boiling point. This allowed us to relate a "distillation cut", such as 95%, to a PMI that would result from such a cut. The estimated PMIs after 95%, 97%, and 99% cuts are shown in Figure 26 as a function of the fuel original PMI. Clear linear relationships between the estimated and original PMI are

observed, with the largest reductions expected for the highest PMI fuels. However, it should be noted again that this estimation method is based on the highly specific chromatographic separation used in D6730 and D8369 methods. Real-life distillation processes are much less chemically resolved, with compounds distilling not as perfectly separated chromatographic peaks, which is the case in the chromatographic methods. Thus, real-life distillation cuts will result in PMIs that are different from those estimated using a DHA method, as was done here.

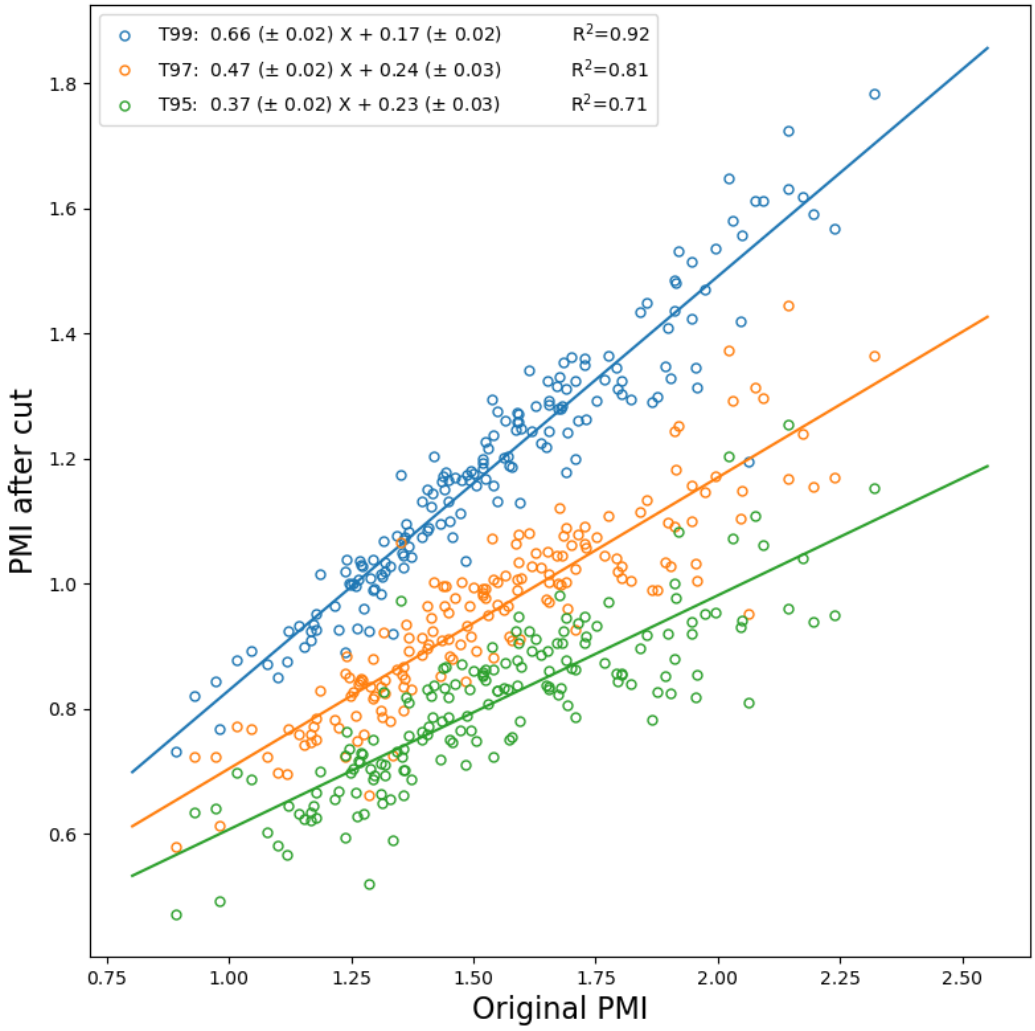

Figure 26. Estimated PMI after a distillation cut (95%: T95, 97%: T97, and 99%: T99) as a function of the original PMI for the summer D6730 dataset.

5.2. Winter dataset

5.2.1. Overview of the winter dataset

The number of fuel samples per grade and PADD region in the winter dataset are given in Table 4. The D8369 and D7096 have 85 samples each, while the D6730 data are missing one sample (Gage # 15735, Exxon regular unleaded, Philadelphia, PA). The D86 is missing three samples Gage # 15689, 15716, and 15721. There were, however, four extra samples in the D86 data that matched descriptions of samples with Gage # 16012, 16021, 16022, and 16036. Since none of these were in the winter datasets of any other measurement methods (and their descriptions did not match any of the three missing samples), they were discarded.

As with the summer dataset, regular unleaded fuels were most represented (9 – 13 per PADD), while there were only four E15 samples (two in PADD 2 and one each in PADD 1 and 3). PADD 1 was the most represented with 22 samples, while PADD 2 through 5 had 20, 18, 10, and 15 samples, respectively.

Table 4. Number of samples for each of the measurement method per fuel grade and PADD.

PADD	grade	D6730	D8369	D86	D7096
1	E15	1	1	1	1
	Premium Unleaded	8	8	8	8
	Regular Unleaded	12	13	13	13
2	E15	2	2	2	2
	Premium Unleaded	5	5	4	5
	Regular Unleaded	13	13	12	13
3	E15	1	1	1	1
	Premium Unleaded	5	5	5	5
	Regular Unleaded	12	12	12	12
4	Premium Unleaded	4	4	4	4
	Regular Unleaded	6	6	6	6
5	Premium Unleaded	6	6	6	6
	Regular Unleaded	9	9	8	9
Total		*84*	*85*	*82*	*85*

5.2.2. Chemical composition data intercomparison for the winter dataset

The range of reported RI for the winter survey was somewhat larger than that for the summer, with the maximum RI reported by the D6730 method being 2052.3, while that for the D8369 was to 1600. The named (both identified and generic) compounds had generally lower RI than the maximum RI reported, indicating the difficulty identifying heavier compounds that tend to have higher RIs.

Table 5. Maximum retention indexes for the entire winter dataset as reported for D6730 and D8369 analyses.

	D6730		D8369	
	Named	All	Named	All
mean	1792.3	1867.3	1305.2	1306.6
std	156.7	183.0	79.2	78.8
min	1409.4	1413.1	1128.1	1128.1
25%	1617.5	1708.9	1274.1	1274.1
50%	1832.1	1952.4	1300.0	1300.0
75%	1951.8	2027.4	1351.3	1351.3
max	1953.6	2052.3	1600.0	1600.0

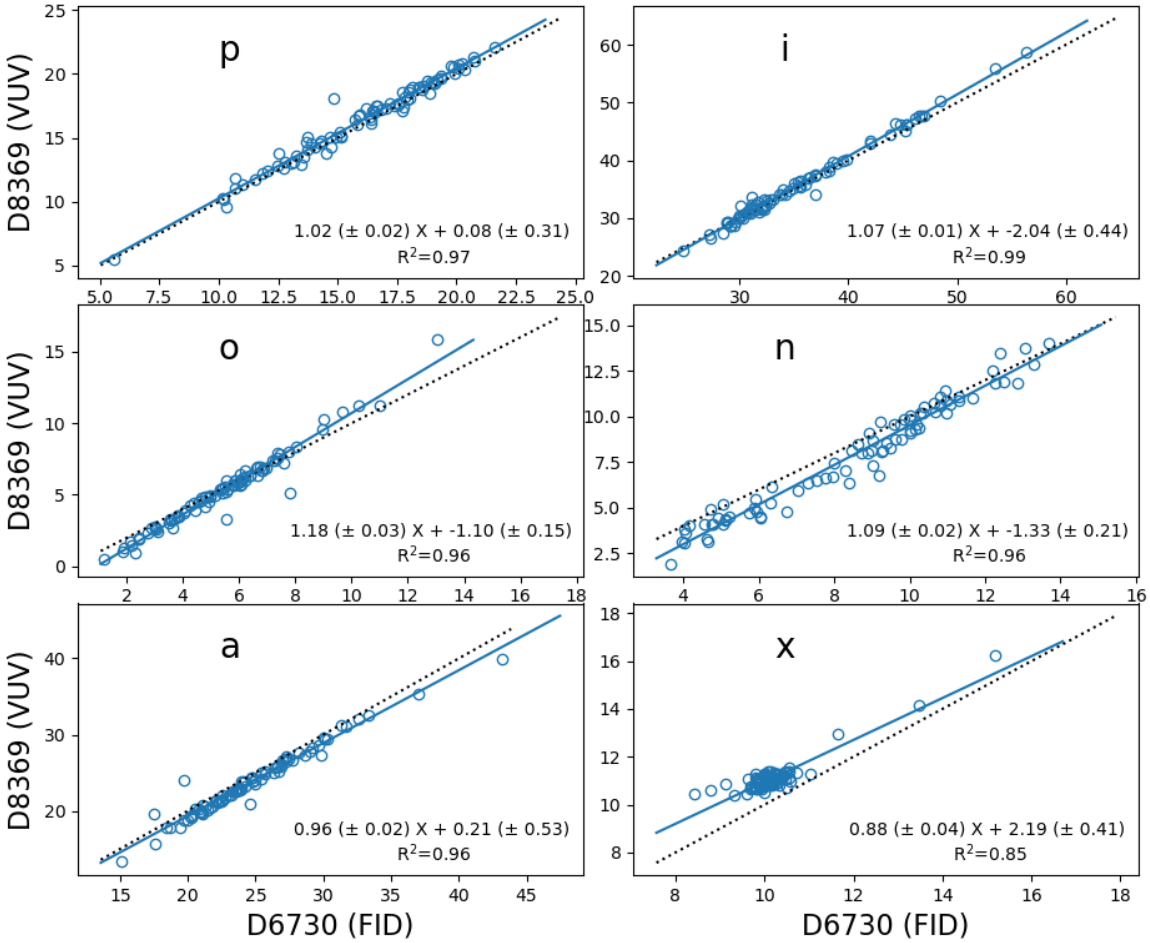

Figure 27. Comparison of total weight % of individual PIONAX groups reported by the D6730 and D8369 methods for the winter dataset. The dotted line shows the 1:1 relationship, while the solid blue line shows the regression line with the regression equation and the corresponding R^2 of the fit given in the lower right corner of each subplot.

The inter-comparison of the FID (D6730) and VUV (D8369) methods shown in Figure 27 is similar to that of the summer dataset, with paraffins, isoparaffins, and aromatics being close to 1:1 and R^2 of a linear fit between the two methods being 0.96. The average ratios of per-sample total weight contributions of PIONAX classes are shown in Figure 28. Overall, the relative performance of the two methods is similar to that seen within the summer dataset, though for oxygenates the FID/VUV ratio was lower in winter. Similar to the summer dataset, the agreement between the two methods is poorer for olefins and naphthenes, with 15%-28% variability around the mean.

A per-compound comparison of the two methods also shows features that are similar to the summer-intercomparison (Figure 29). Generally, the agreement is good for concentrations higher than approximately 1% by weight, while for low concentrations the two methods report values that differ by up to two orders of magnitude. Likewise, the total PMIs reported by the two methods correlate with each other, with the FID method reporting on average 28-29% higher PMI values (Figure 30).

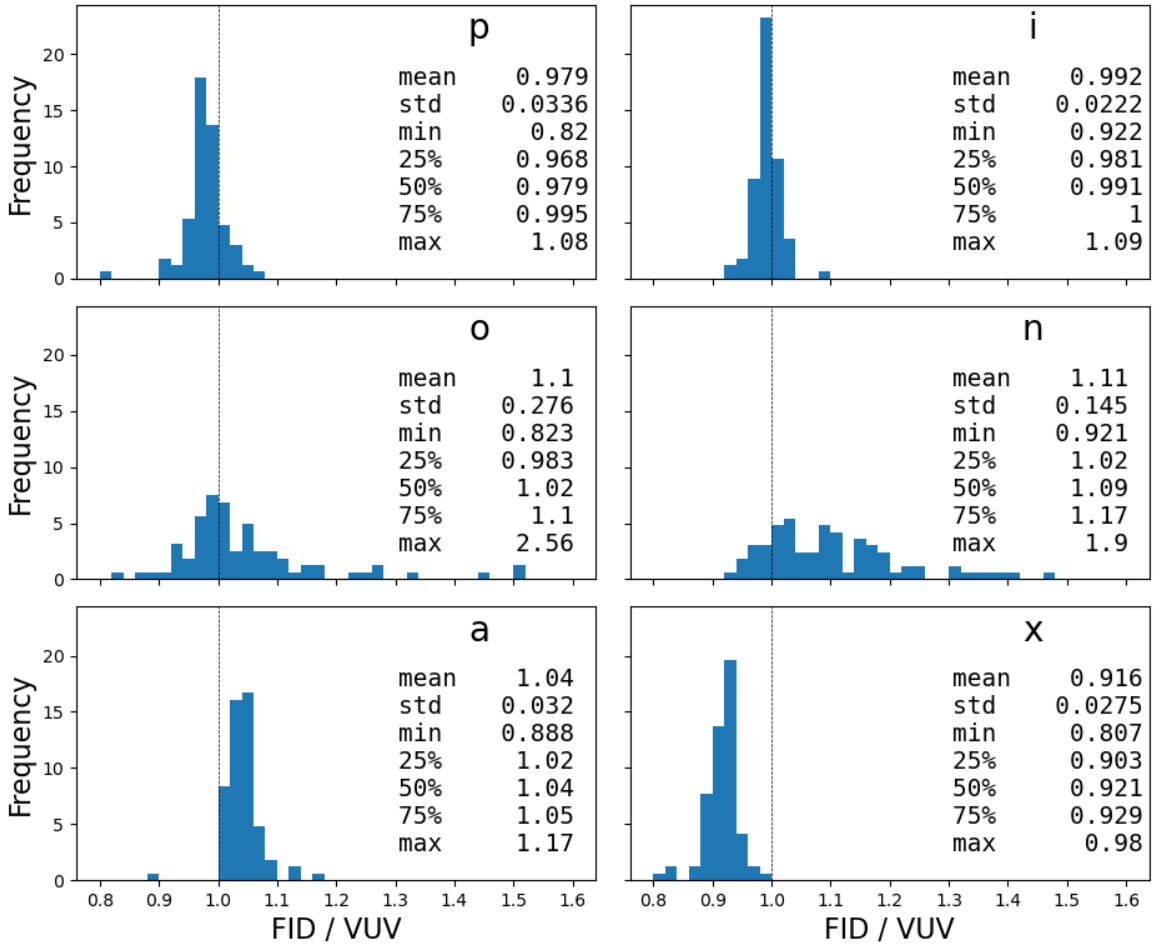

Figure 28. Frequency distributions of per-fuel ratios of total PIONAX group contributions reported by the FID (D6730) to the VUV (D8369) methods for the winter dataset. The descriptive statistics of the ratios are also provided.

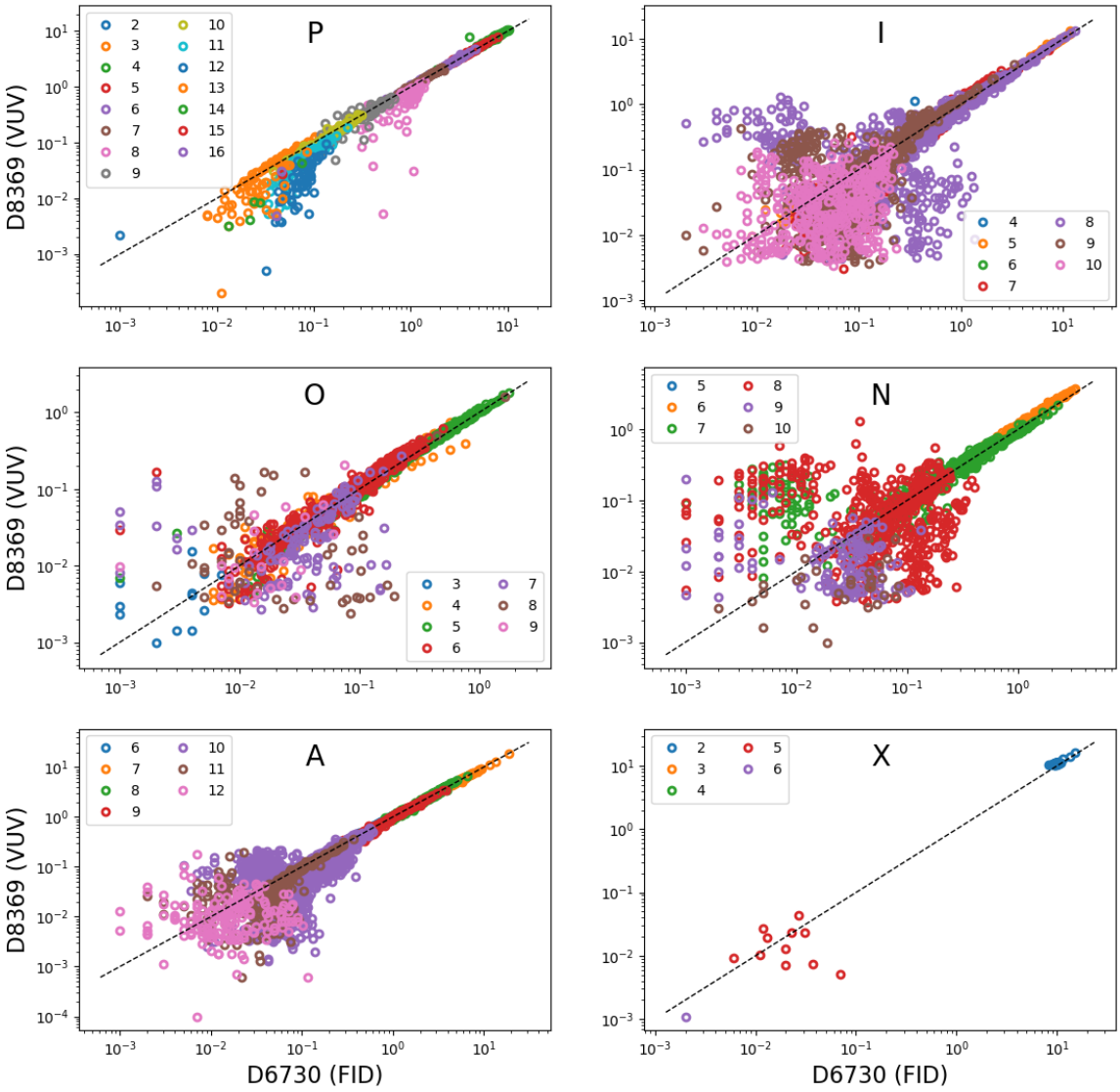

Figure 29. Per compound comparison within individual PIONAX groups reported by the D6730 and D8369 methods for the winter dataset. Compounds with the same carbon number are indicated by the same color (carbon numbers are given in the legends). The dotted line shows the 1:1 relationship.

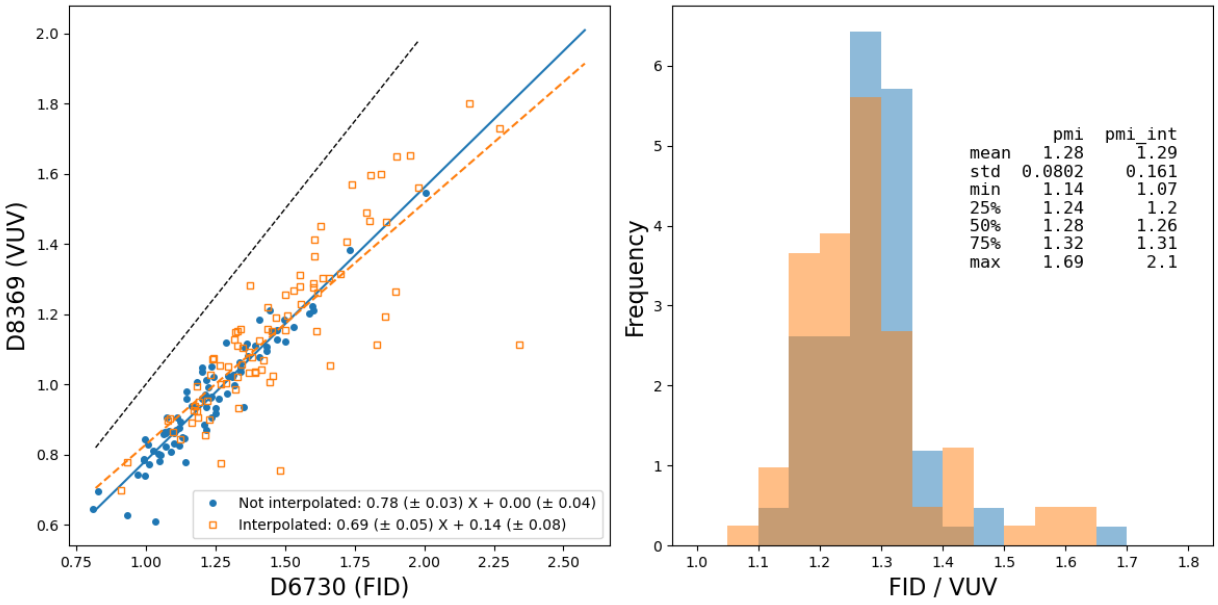

Figure 30. Left: Comparison of PMIs reported by D6730 (FID) and D8369 (VUV) methods for the winter dataset ("Not interpolated" – PMI data as reported, "Interpolated" – including interpolated PMI data). A broken black line represents 1:1 relationship. Right: Histogram of FID/VUV ratios (the insert provides descriptive statistics of the ratios; "pmi" – PMI data as reported, "pmi_int" – including interpolated PMI data).

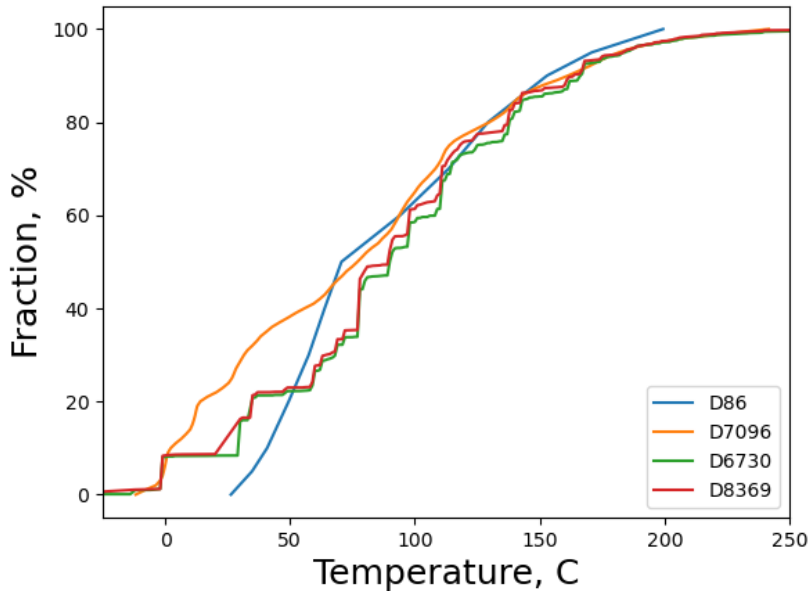

Figure 31. Comparison of the average distillation profiles measured with the D86 and D7096 methods and derived from the D6730 and D8369 data for the winter dataset.

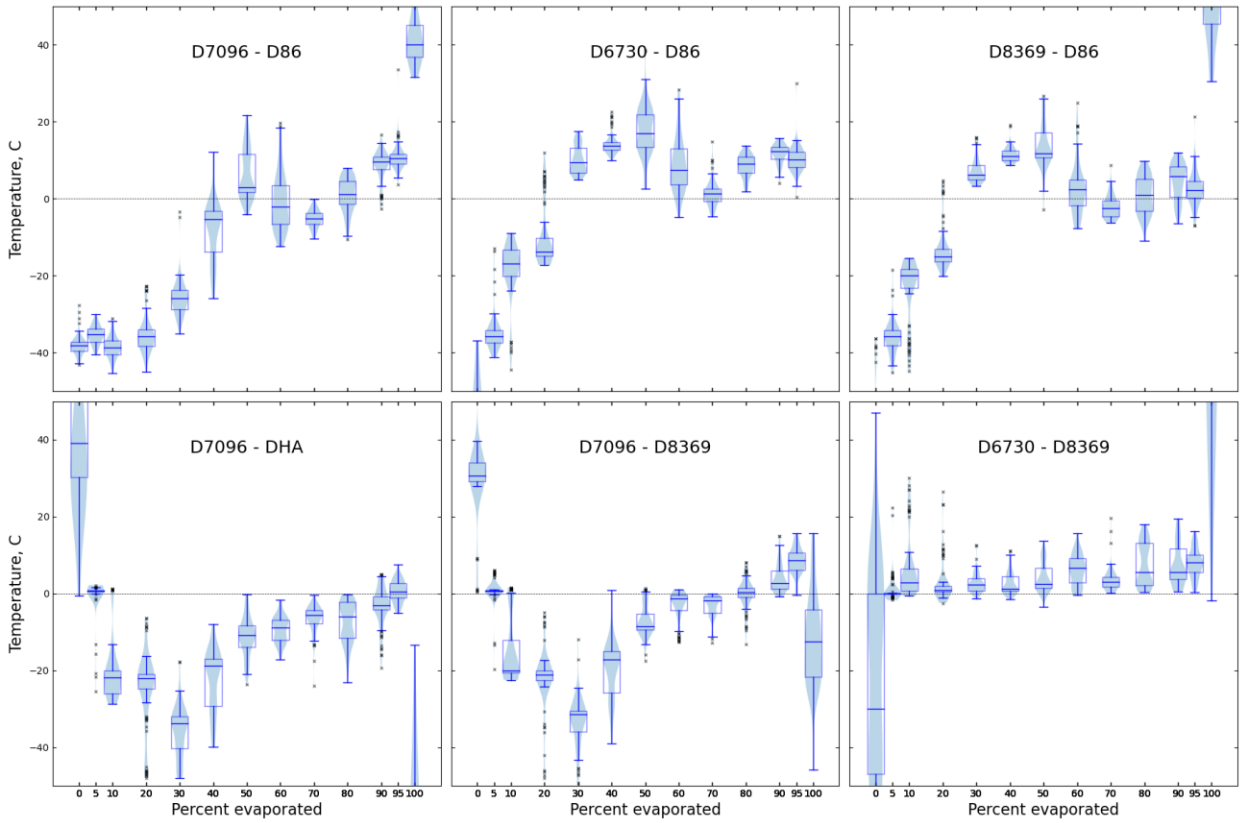

Figure 32. Violine plots and boxplots of per-sample differences in reported temperatures required to evaporate a certain fraction of fuel for the winter dataset.

5.2.3. Winter distillation profile intercomparison

Average distillation profiles for the winter dataset as reported by the D86 and D7096 methods and those derived from the D6730 and D8369 methods, are shown in Figure 31. Similar to the summer dataset, all methods agree fairly well with each other between approximately 80ºC and 200ºC. The differences between the methods are also similar to those described in Section 5.1.3. Likewise, the per-sample comparison of distillation profiles is similar to that of the summer dataset (Figure 32).

5.2.4. Group contributions to weight and PMI

The average per group composition for the summer dataset is shown in Figure 33 and summarized in Table 6. The results are generally similar to those for the summer dataset, though there is a larger content of paraffins in the winter fuels (16% vs. 10.9%), especially C4 paraffin. Butane is excluded from summer gasoline to meet RVP limits, but included in winter gasoline. Compensating for the higher paraffin content, the other components are less abundant, except for isoparaffins: the aromatic content is about 2% by absolute weight contribution lower, while the other components are about 1 weight % lower. The weight content of unknown compounds is larger in winter, though still very small: 0.23% vs. 0.086% in summer.

The aromatic contribution to the overall PMI is also practically the same as in summer, contributing approximately 88% according to both the FID and VUV methods. The contribution of other components to the overall PMI is also similar to that in summer. The variability of the group weight and PMI contributions is shown in Figure 34 using the FID data. Figure 35 shows a more detailed view of the contribution of aromatic compounds by carbon numbers to the weight and PMI for the D6730 winter dataset. The picture is generally similar to that in summer, with C10 aromatics contributing on average most to the PMI. However, the contribution of C13 and C14 aromatics is more pronounced in winter.

Table 6. Average contributions of PIONAX groups to weight and PMI for winter dataset. "PMI interpolated" includes interpolated PMI values for generic and unknown compounds.

	Weight %		PMI		PMI interpolated	
	FID	VUV	FID	VUV	FID	VUV
group						
p	16.0	16.4	0.02	0.02	0.02	0.02
i	35.1	35.5	0.06	0.05	0.08	0.06
o	5.5	5.4	0.01	0.01	0.01	0.01
n	8.4	7.9	0.04	0.02	0.05	0.04
a	24.7	23.7	1.10	0.86	1.29	1.02
x	10.2	11.1	0.01	0.01	0.01	0.01
u	0.1	0.0	0.00	0.00	0.01	0.00

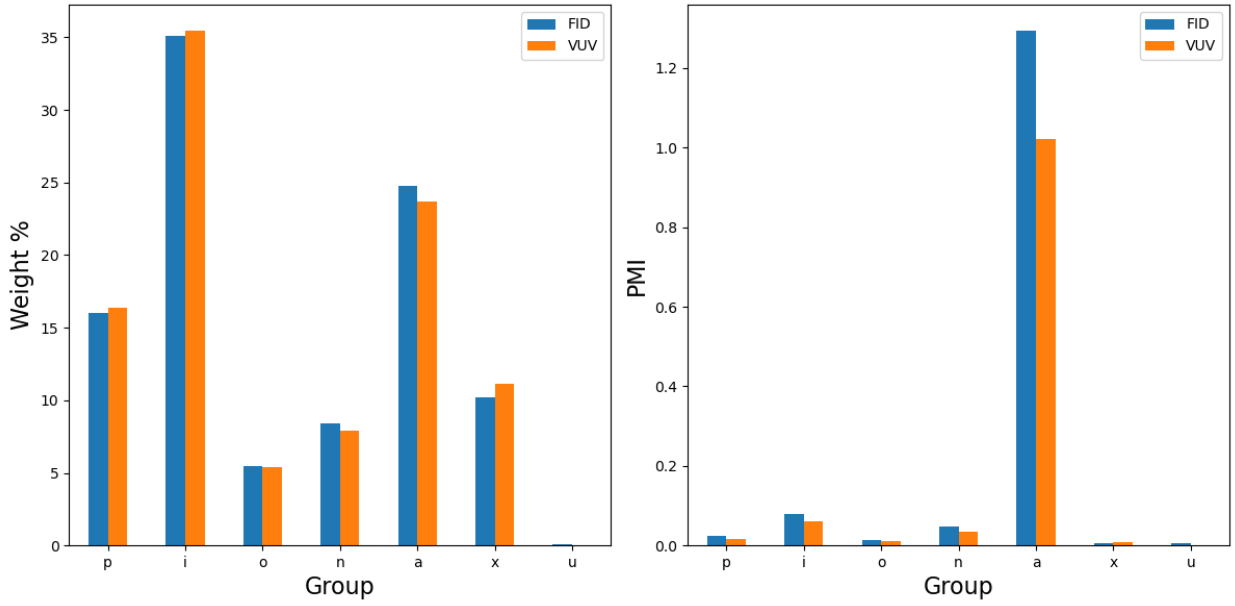

Figure 33. Average contributions of the PIONAX groups to fuel weight composition and PMI for the winter dataset.

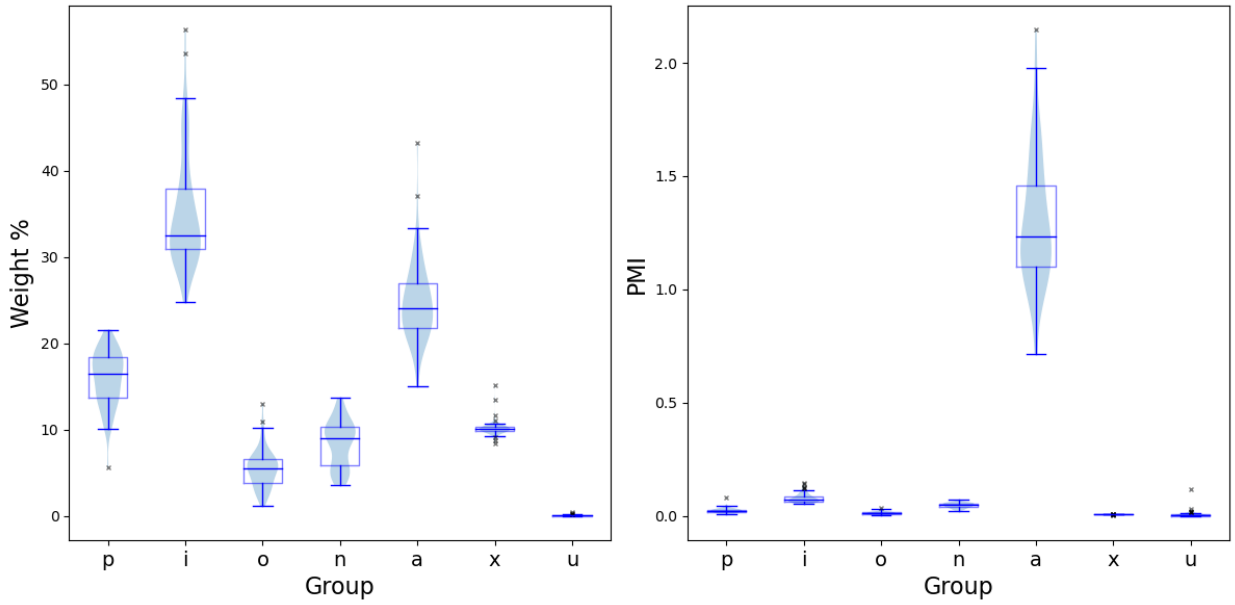

Figure 34. Compound group contributions to weight and PMI for the winter dataset based on the FID (D6730) data.

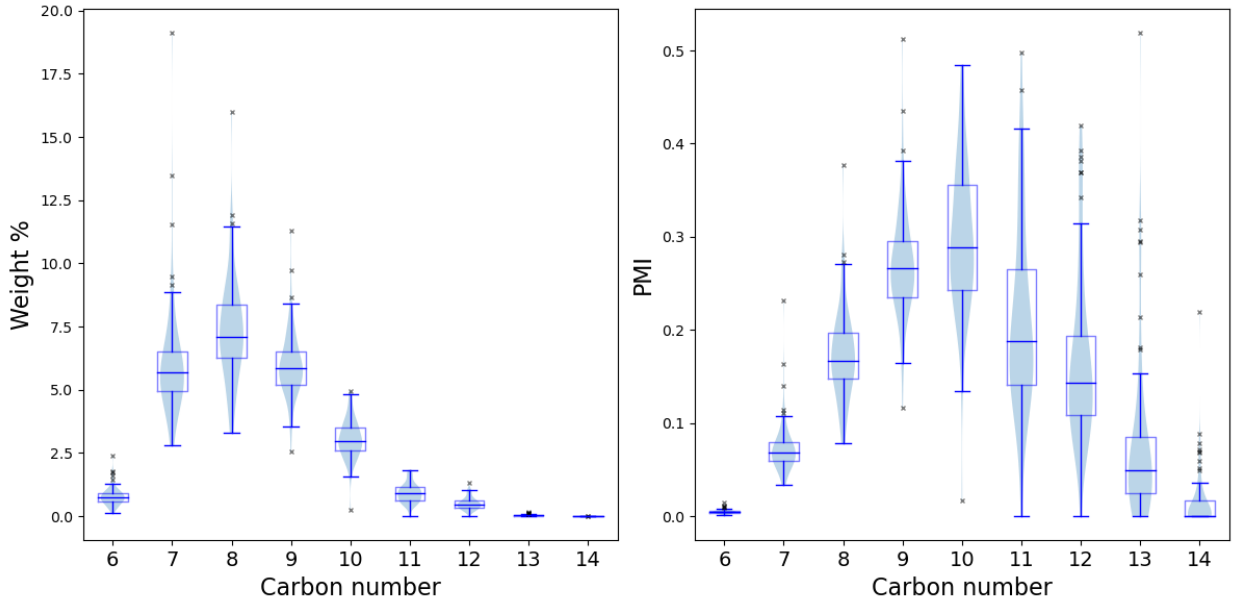

Figure 35. Contribution of aromatic compounds to weight and PMI for the D6730 winter dataset. PMI values were calculated including interpolated values for generic compounds.

5.2.5. Comparison of winter gasolines across grades and PADD regions

The average winter composition of each of the three fuel categories per PIONAX group and carbon number is shown in Figure 36. The overall breakdown per PIONAX group and carbon

number is similar to that of the summer fuels. However, a notable difference is that all grades of winter fuels have butane as the most abundant paraffin, unlike pentane in the summer fuels.

Similar to the summer datasets, the statistical significance of differences between different grades and PADD regions of winter fuels was checked using the Wilcoxon signed rank test. As in the summer fuels, the premium unleaded gasoline has a much higher content of C8 isoparaffins (Figure 36). This leads to statistically significant differences of this fuel grade in this compound class, as well as several other classes (Figure 37). Unlike the summer fuels, there are statistically significant differences in ethanol content between all three fuel grades, though the difference between the regular and premium gasolines is very small: 10% vs. 9.8%, respectively.

There are a few statistically significant differences between PADD regions in winter fuel composition of premium unleaded (Figure 38) and regular unleaded (Figure 39) gasolines. No statistical analysis of E15 fuels was possible due to their very low number. Generally, there are fewer and weaker differences in the winter fuels, probably because of the smaller number of samples in this dataset.

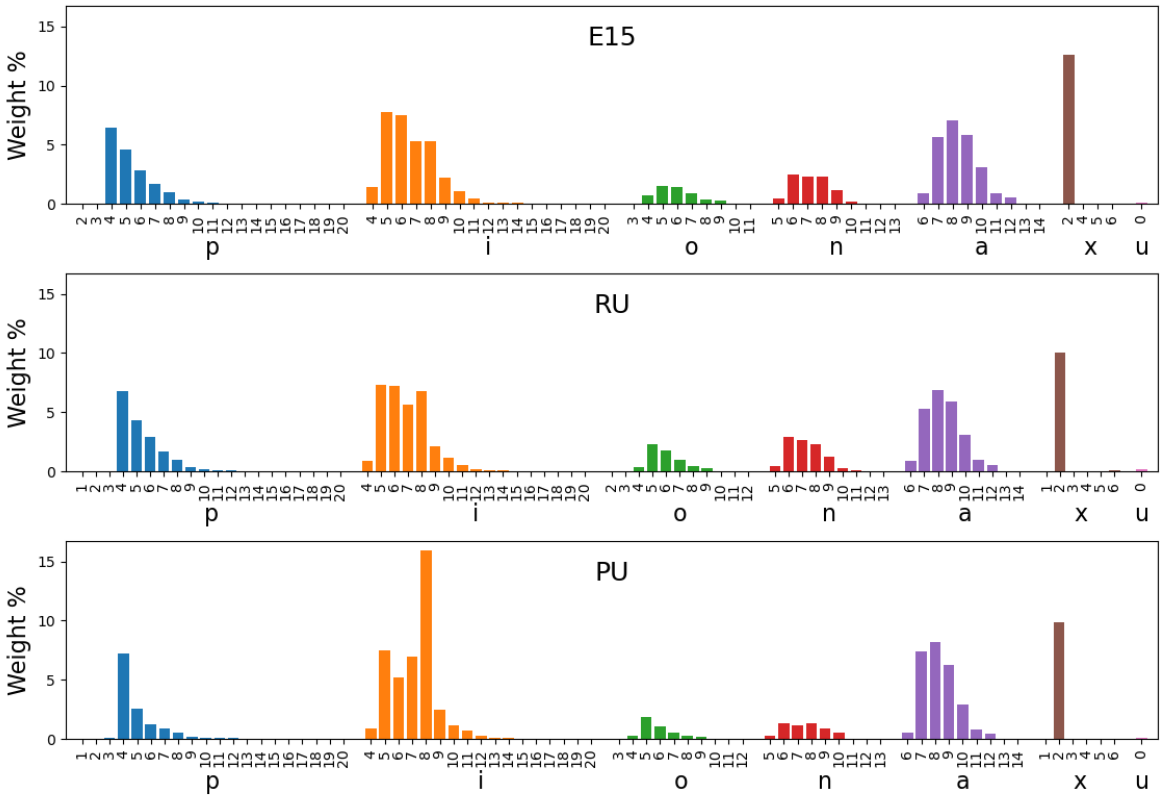

Figure 36. Average composition of the three gasoline grades per PIONAX group and carbon number based on D6730 winter data (RU: regular unleaded, PU: premium unleaded).

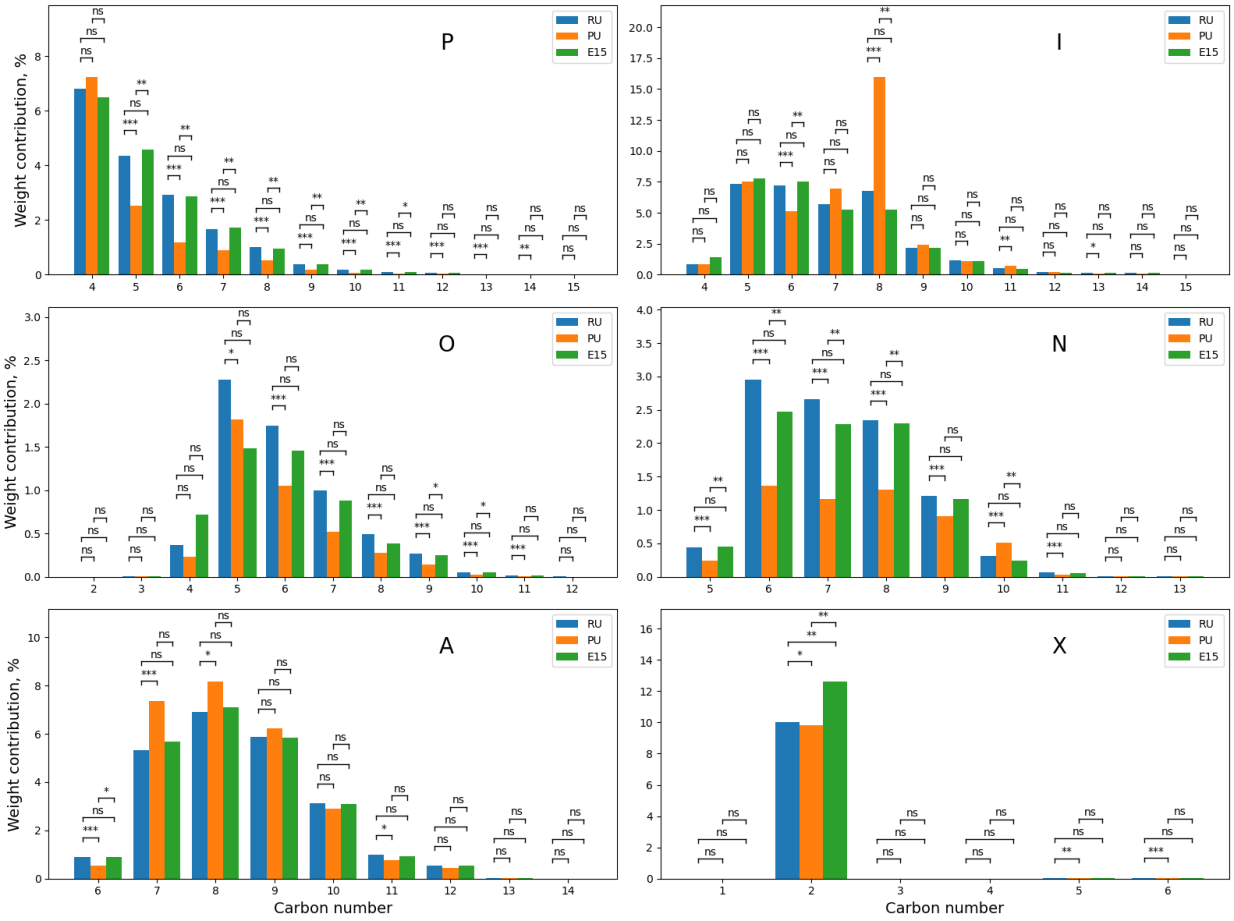

Figure 37. Statistical comparison of chemical composition of the three fuel grades using D6730 winter data (RU: regular unleaded, PU: premium unleaded). Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

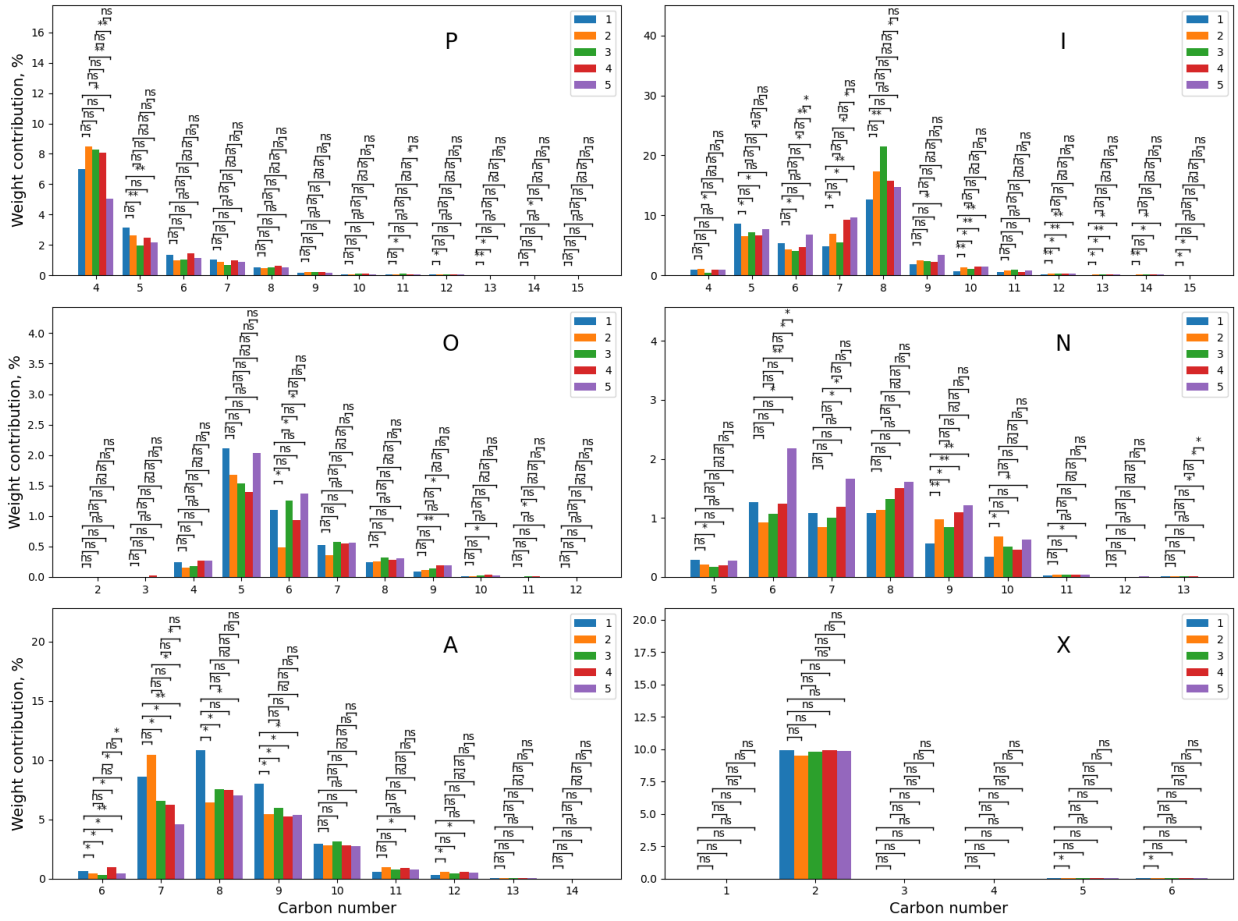

Figure 38. Statistical comparison of chemical composition of winter premium unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

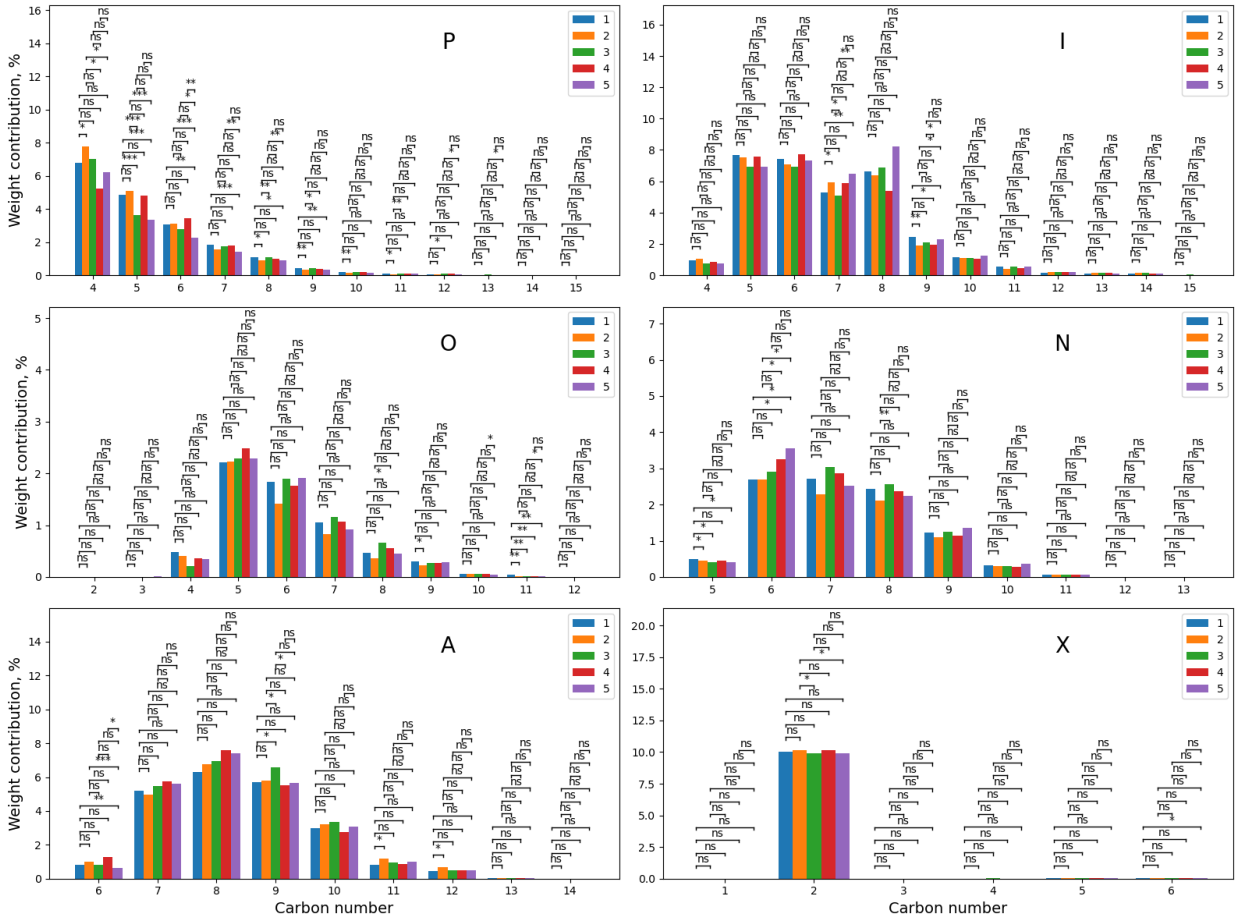

Figure 39. Statistical comparison of chemical composition of winter regular unleaded gasoline in five PADD regions broken down by PIONAX groups. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

5.2.6. Comparison of PMI across grades and PADD for winter dataset

A comparison of PMI values (including interpolated values) between the fuel grades and the five PADD regions for the entire winter D6730 dataset is shown in Figure 40. The Wilcoxon test was applied to assess the statistical significance. Unlike the summer dataset, where the premium unleaded fuels are statistically significantly lower in PMI than the other two grades, there are no significant differences between the grades in the winter dataset. Statistically significant differences in PMI are observed between PADD 1 and 2, and PADD 2 and 5, which is similar to the summer dataset, though no significance was observed between PADD 2 and PADD 3 and 4, probably due to the smaller number of samples in the winter dataset. Unlike the summer dataset, no significant differences were observed between any PADD regions for premium unleaded fuels (Figure 41). For regular unleaded gasolines, differences in PMI between PADD 1 and 2, 1 and 3, and 2 and 4 are observed, while for summer only two differences were significant (between PADD 1 and 2, and 2 and 5). No comparison was possible for E15 fuels due to the low number of samples.

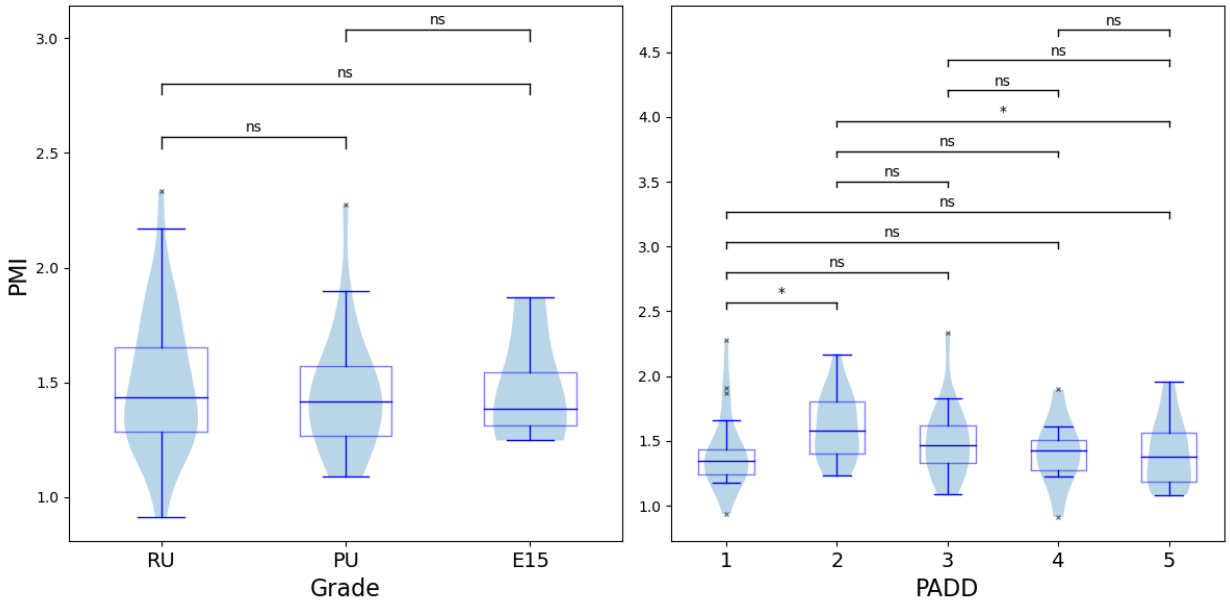

Figure 40. Statistical comparison of PMI of different grades of fuel and fuels in different PADD regions for the entire winter dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

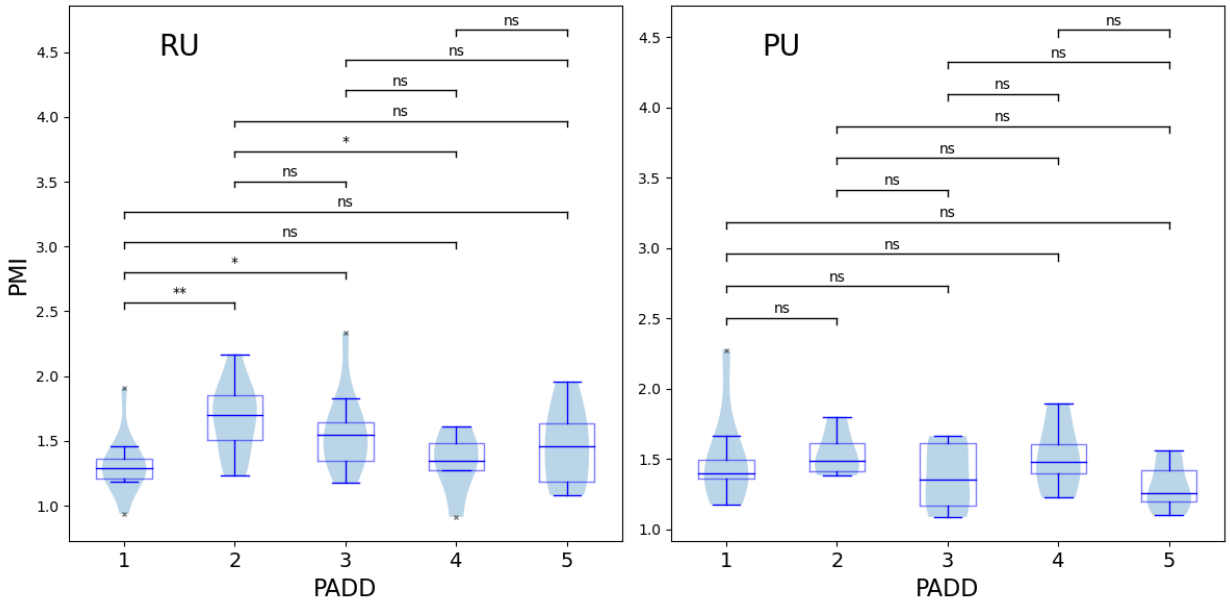

Figure 41. Statistical comparison of PMI between different PADD regions for two fuel grades for the winter dataset based on D6730 analysis and using interpolated values for generic and unknown compounds. The comparison is based on D6730 analysis. Three stars denote p < 0.001, two stars: p < 0.01, one star: p < 0.05, and "ns" for p values above 0.05 (not significant at 5% confidence).

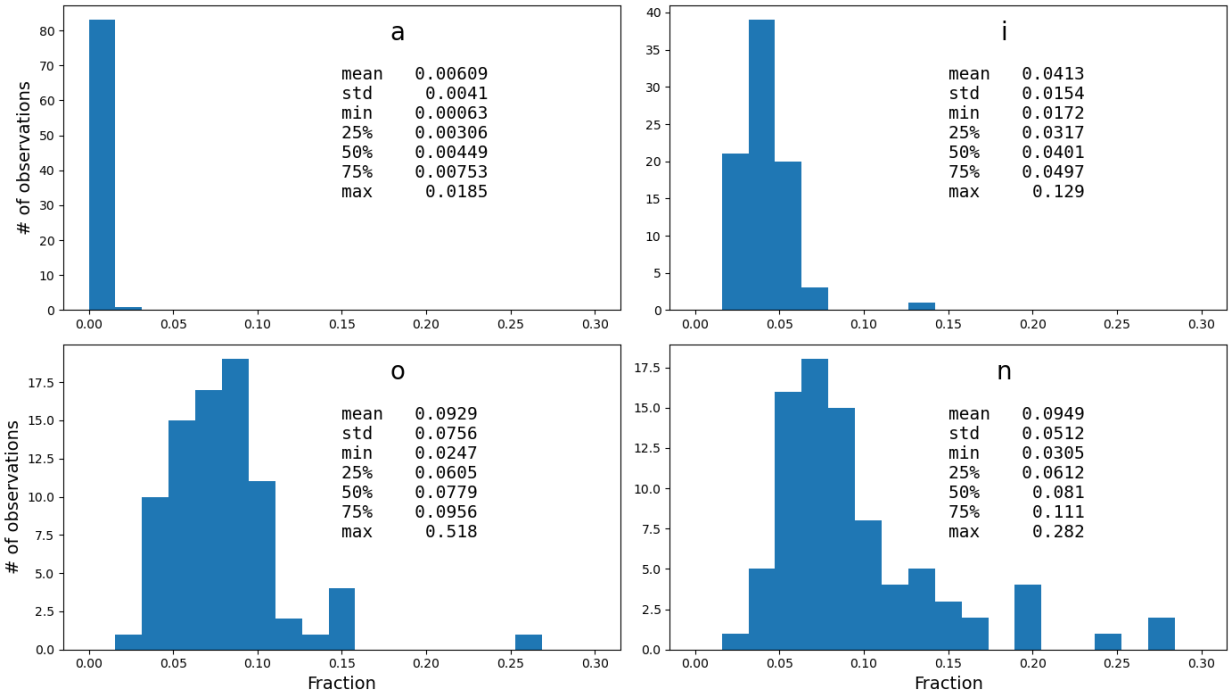

Figure 42. The weight fraction of generic compounds within aromatic ('a'), isoparaffin ('i'), olefin ('o'), and naphthene ('n') groups for the winter D6730 dataset (all grades and PADD regions).

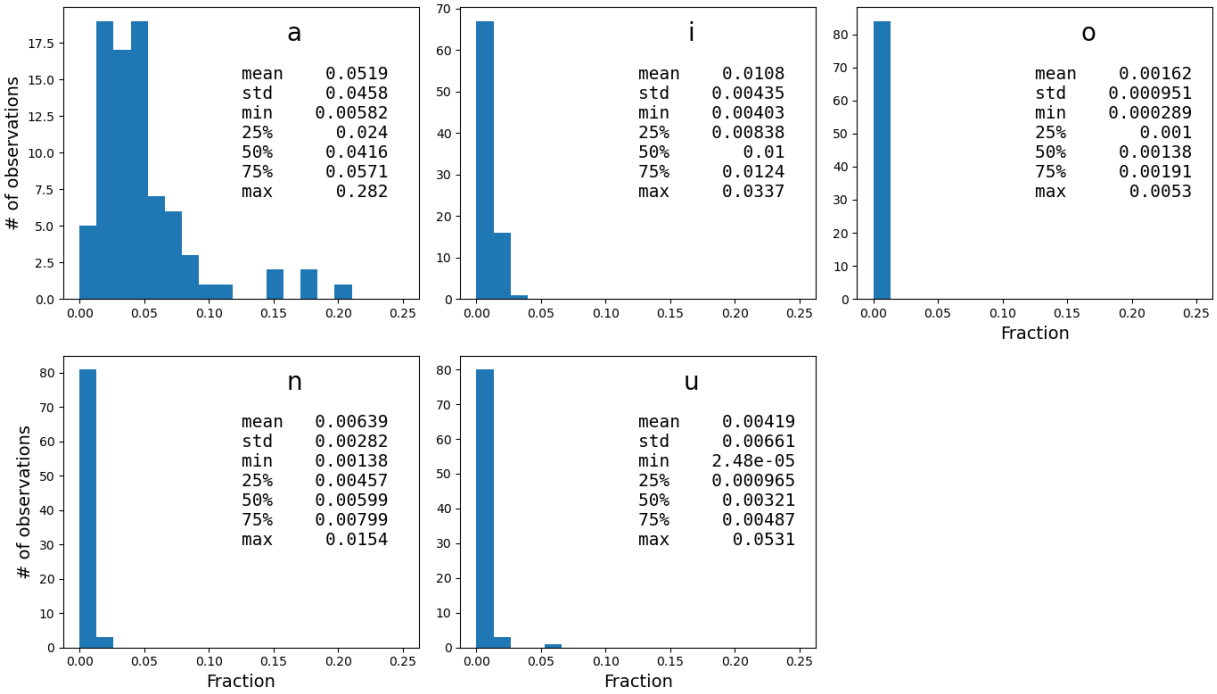

Figure 43. PMI contribution of generic and unidentified compounds to the total fuel PMI based on the D6730 winter dataset (all grades and PADD regions). 'a': aromatics, 'i': isoparaffins, 'o': olefins, 'n': naphthenes, 'u': unidentified.

5.2.7. Contributions of generic compounds to weight and PMI in winter fuels

The contribution of generic compounds to the total weight contribution by PIONAX groups for the D6730 winter dataset is broadly similar to that observed in the summer dataset (Figure 42). The contribution of generics is smaller in winter than in summer for aromatics (0.6% vs. 1.1%) and naphthenes (9.5% vs. 12.8%), while their contribution to isoparraffins and olefins is approximately the same as in summer. The contribution of generics to the total fuel PMI is practically identical to that for the summer dataset (Figure 43), with the largest contribution the generic aromatics compounds of about 5% ± 3% of the total fuel PMI.

5.2.8. Winter distillation profiles

The average distillation profiles determined by different methods for the winter dataset are shown in Figure 44. Similar to the summer dataset, the biggest differences are observed for the lowest and highest temperatures, with the FID method (D6730) extending most to the higher temperatures due to its larger chromatographic range. Similar to the summer dataset, the variability among individual distillation profiles is fairly small.

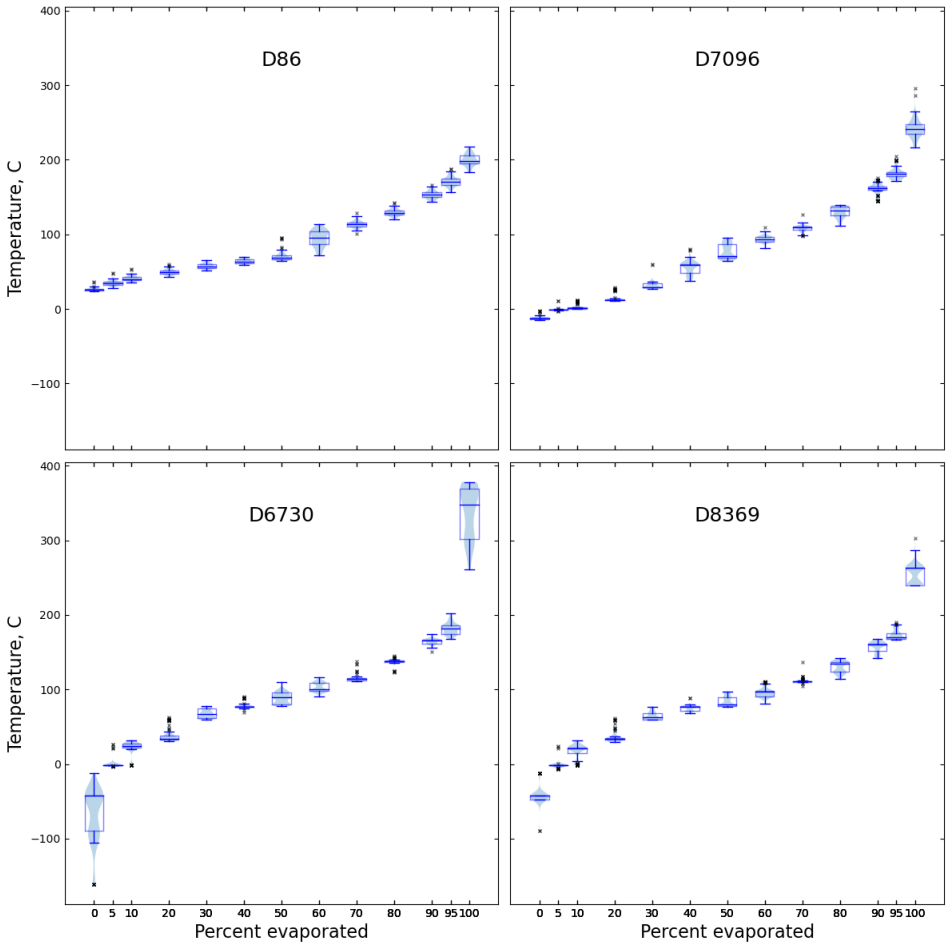

Figure 44. Violin and boxplots of distillation profiles (temperatures required to evaporate a certain fraction of fuel) for winter dataset.

5.2.9. Hypothetical effect of distillation cuts on winter fuel PMI

Similar to the summer dataset, the winter dataset exhibits a linear correlation between the fuel PMI and temperatures at which certain fractions of fuel are evaporated, such as 95%, 97%, and 99% (Figure 45). The correlations are less strong due to the lower number of samples in the winter dataset. Likewise, the expected PMI after a distillation cut linearly correlates with the original PMI of the fuel (Figure 46), with the regression parameters close to those observed in summer.

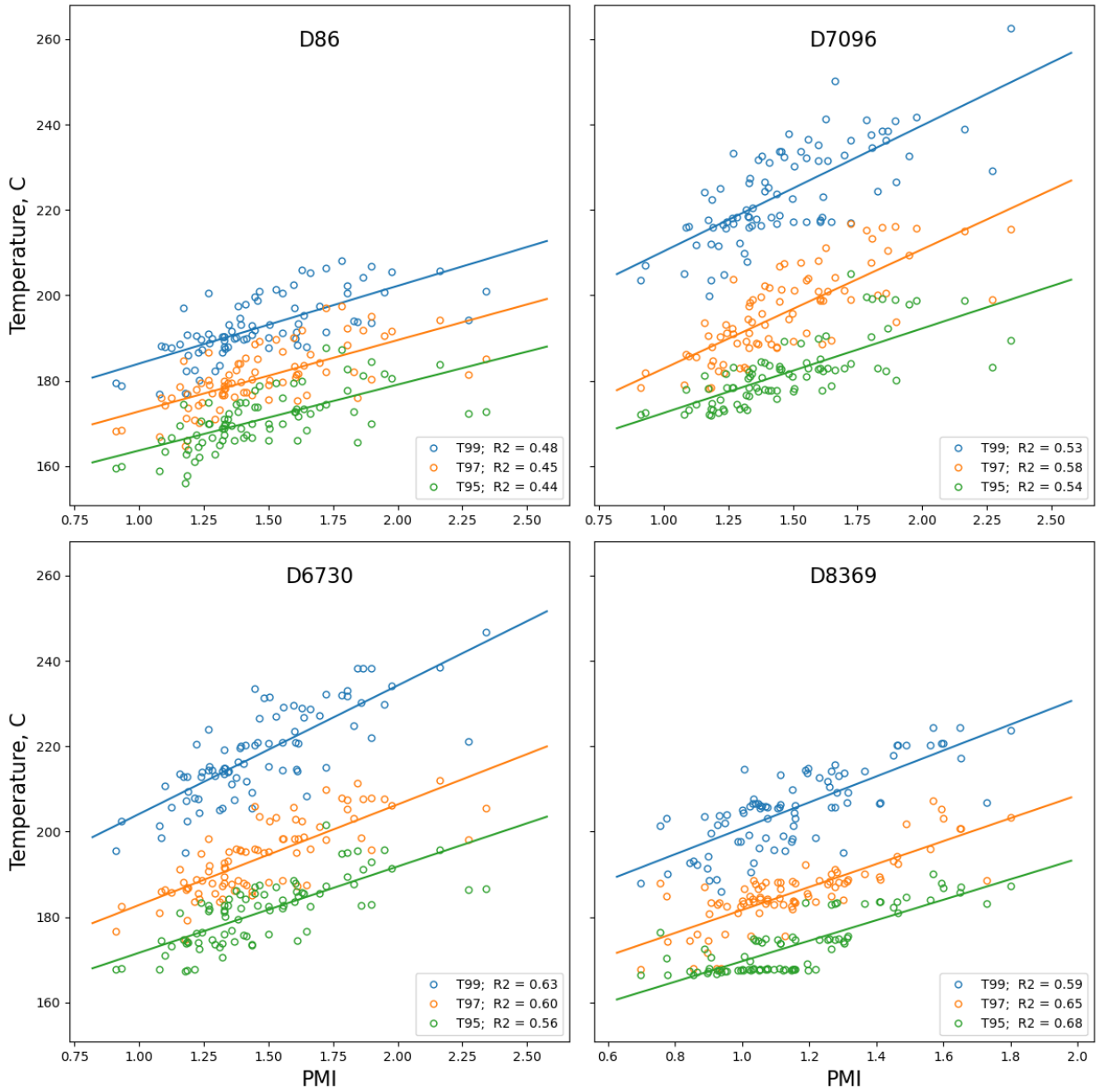

Figure 45. Temperature at which 95% (T95), 97% (T97), and 99% (T99) of fuel is evaporated vs. fuel PMI for the winter dataset for the four measurement methods. The D8369 subplot uses PMI values derived using the D8369 method; the other plots use PMI values derived using the D6730 method.

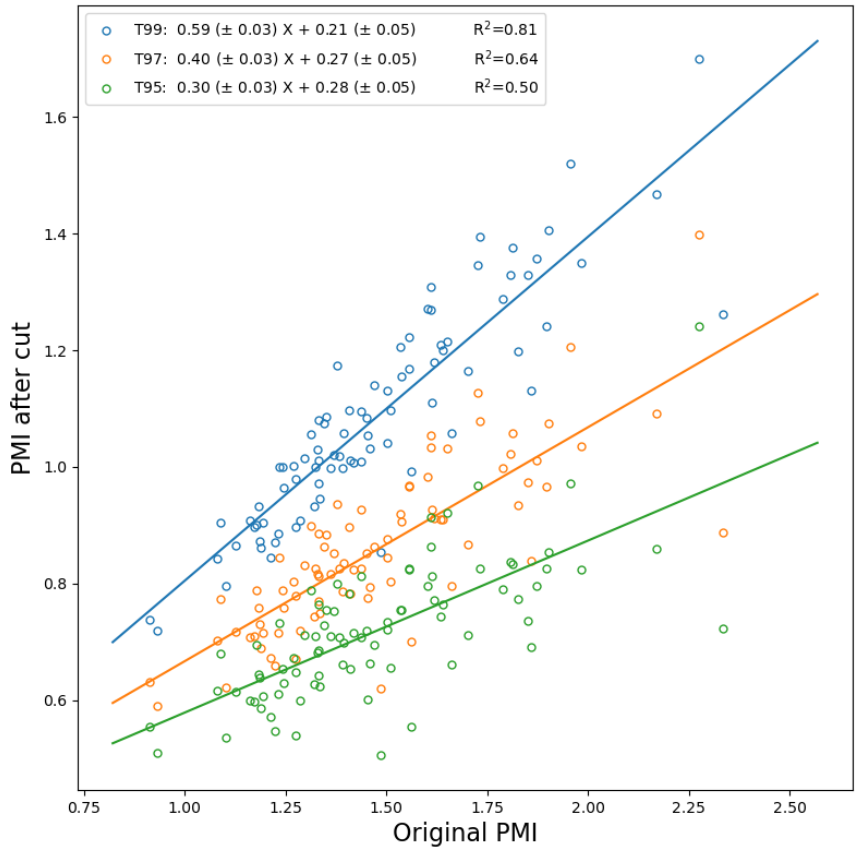

Figure 46. Estimated PMI after a distillation cut (95%: T95, 97%: T97, and 99%: T99) as a function of the original PMI for the winter D6730 dataset.

5.3. Comparison of summer and winter datasets

5.3.1. Comparison of fuel composition and PMI between the two seasons

A per-group and carbon number comparison of fuel composition of the summer and winter datasets is shown in Figure 47. There are several statistically significant differences based on the Wilcoxon test. The most pronounced is the higher content of light paraffins in winter, especially butane. However, octane and decane are somewhat higher in summer. There are statistically significant differences between some heavier paraffins, but their weight content is rather small. C5-C9 olefins are higher in summer, though C4 olefins are higher in winter. There are statistically significant differences in C8-10 isoparaffin content, with the more prominent C8 isoparaffins being higher in summer. C9, 11, 12 naphthenes are higher in summer. C8-11 aromatics are somewhat higher in summer, with the most pronounced difference observed in C8 aromatics. There is a higher ethanol content in summer relative to winter. As will be shown below, the difference is due to the higher ethanol in premium and regular unleaded fuels, while E15 fuels have insignificant differences in ethanol content.

Figure 48, Figure 49, and Figure 50 show a seasonal comparison between premium unleaded, regular unleaded, and E15 fuels, respectively. All fuel grades show differences similar to those observed for the aggregate summer and winter datasets, with the exception that there are fewer statistically significant differences, especially for E15 fuels, due to the smaller number of samples. The ethanol content of all summer fuel grades is slightly higher, though for the E15 fuels this difference is not significant.

There are also statistically significant differences between the seasons in how much individual compound groups contribute to the fuel PMI (Figure 51). C8-10 aromatics contribute more to PMI in summer than in winter, while C13 and C14 aromatics have higher contribution in winter, though their overall contribution relative to C8-10 aromatics is much smaller.

Histograms of fuel PMIs for the two datasets are shown in Figure 52. Overall, summer fuels have a slightly higher average PMI (1.54) than winter fuels (1.48). This difference, however, is not statistically significant at 95% confidence level ($p = 0.062$).

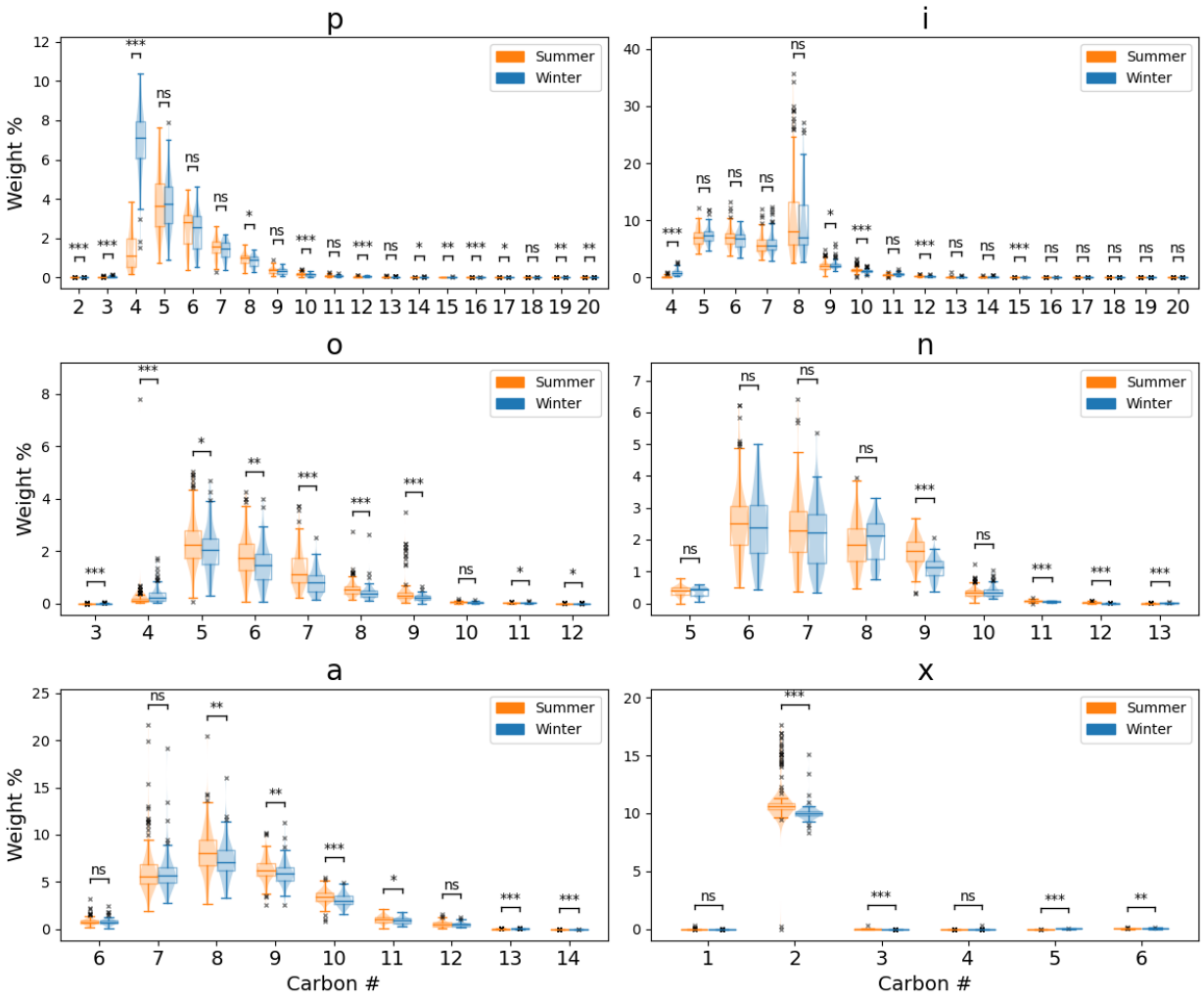

Figure 47. Comparison of weight contributions of PIONAX groups between complete summer and winter datasets.

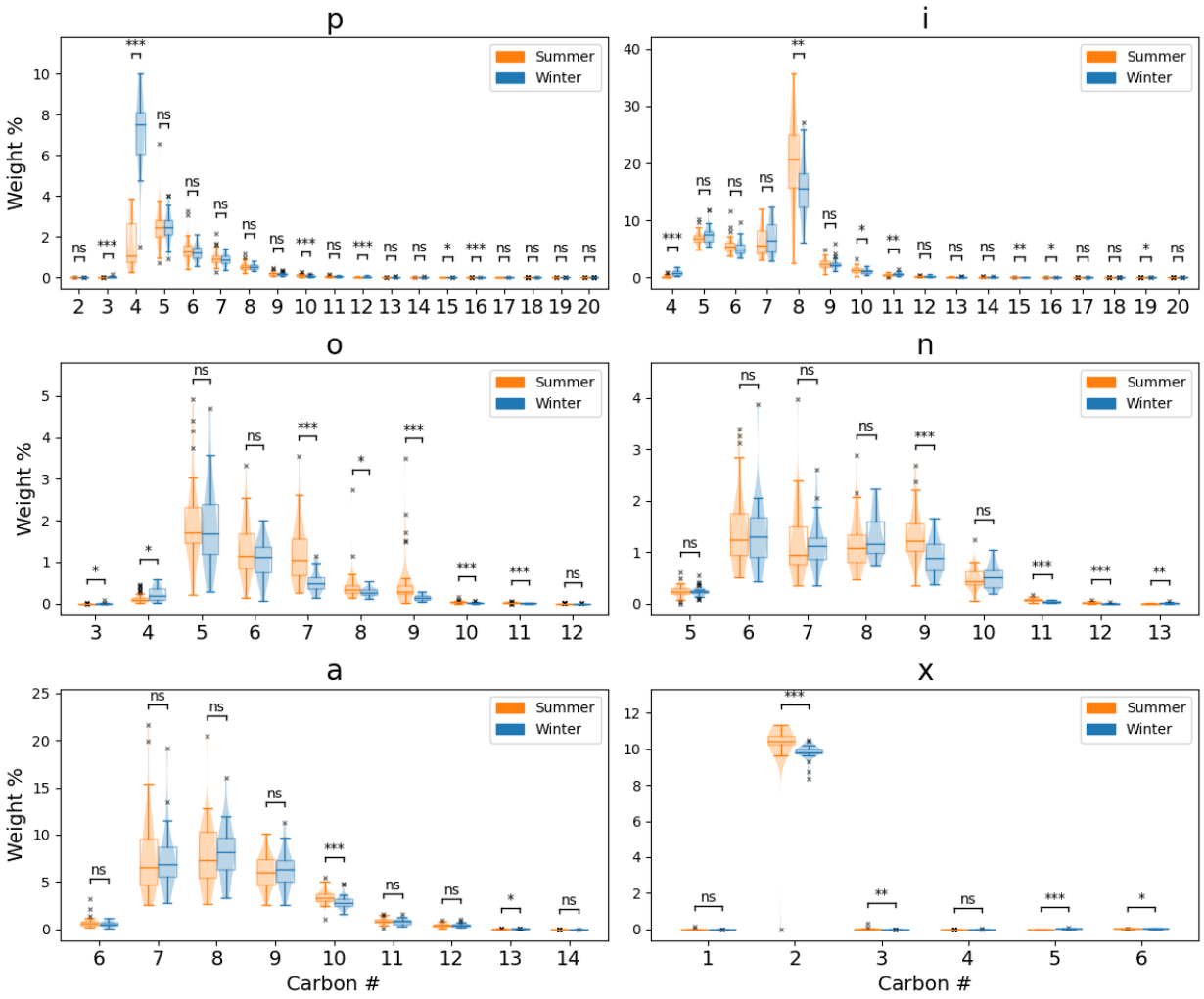

Figure 48. Comparison of weight contributions of PIONAX groups between premium unleaded fuels of summer and winter datasets.

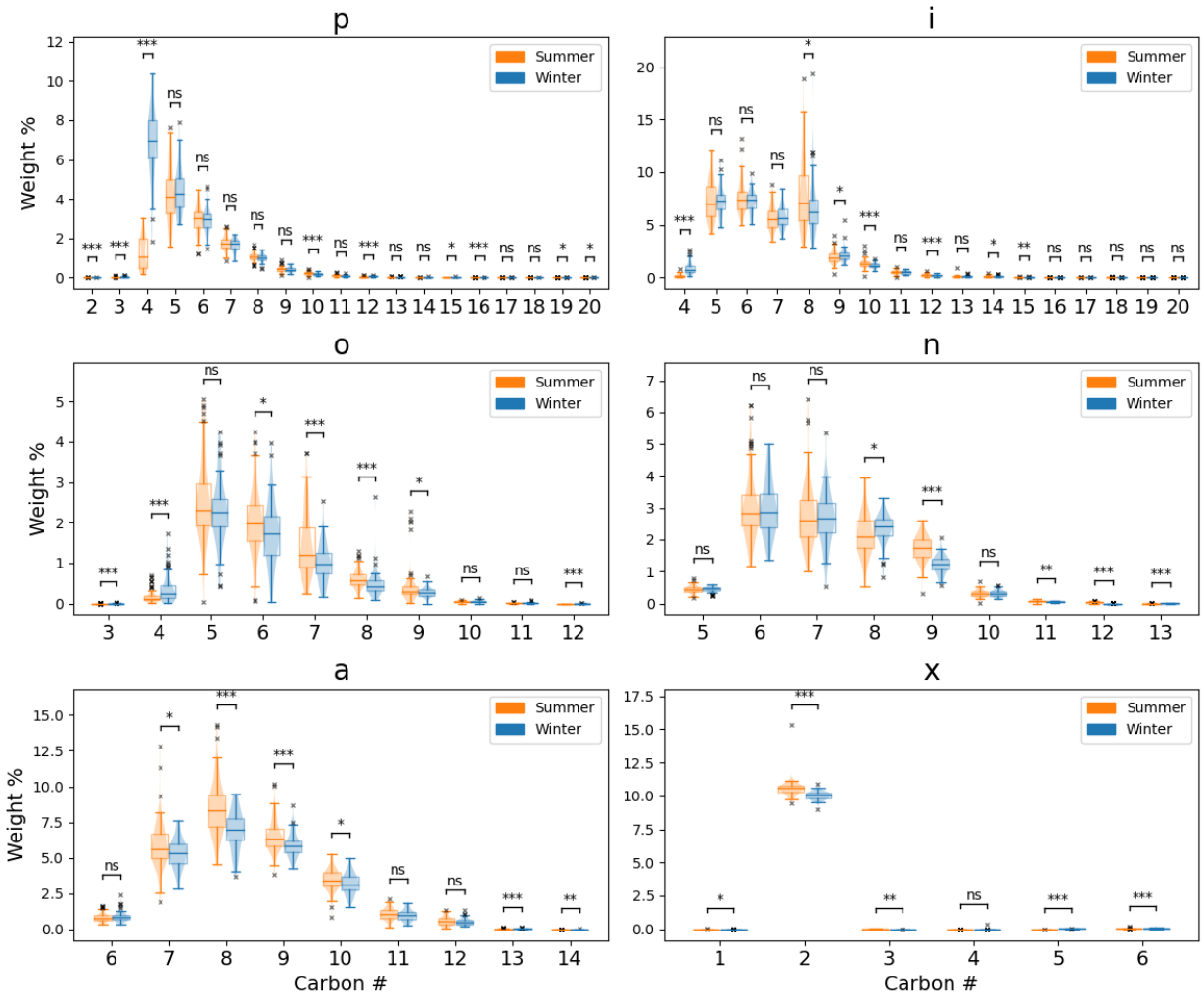

Figure 49. Comparison of weight contributions of PIONAX groups between regular unleaded fuels of summer and winter datasets.

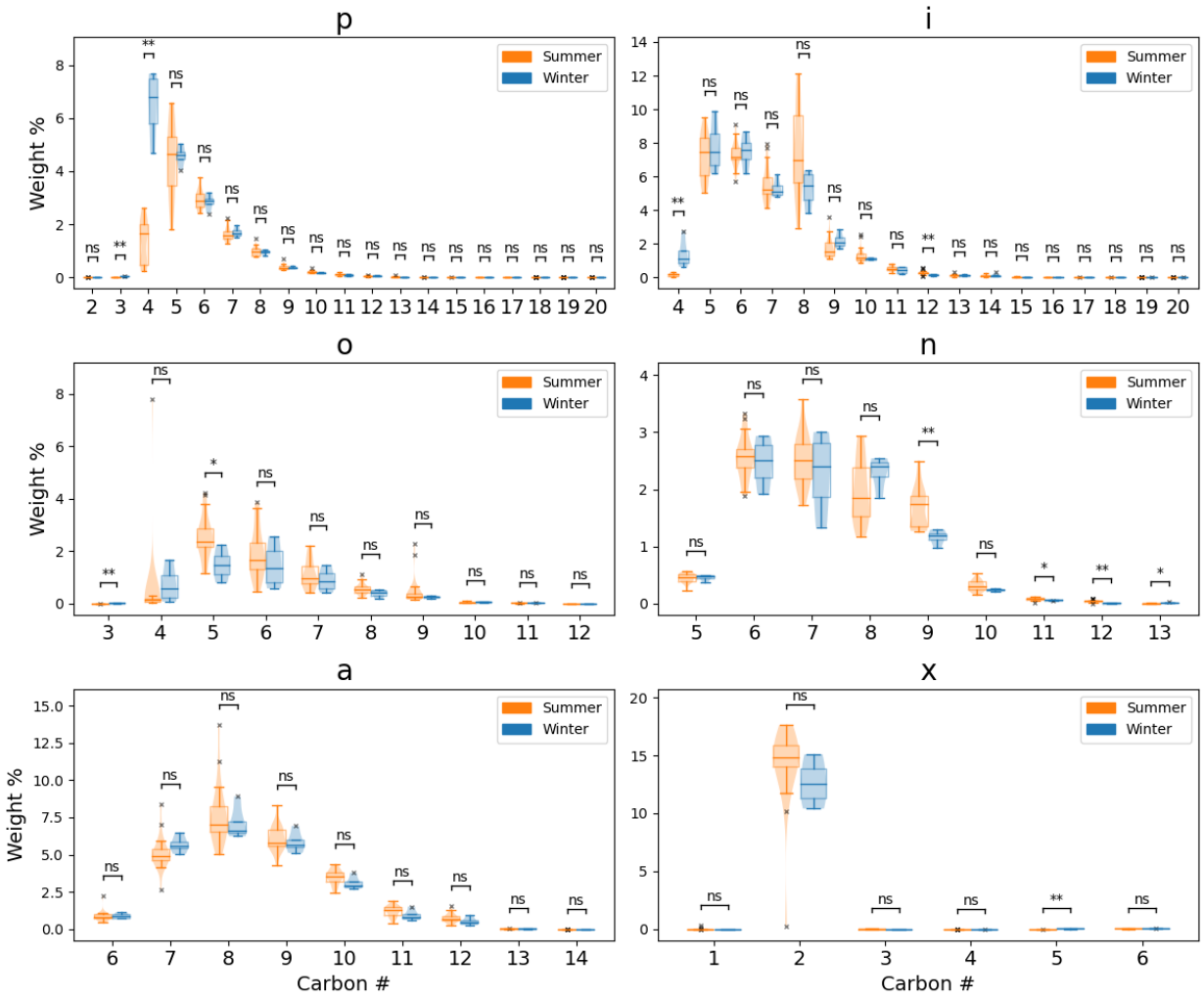

Figure 50. Comparison of weight contributions of PIONAX groups between E15 fuels of summer and winter datasets.

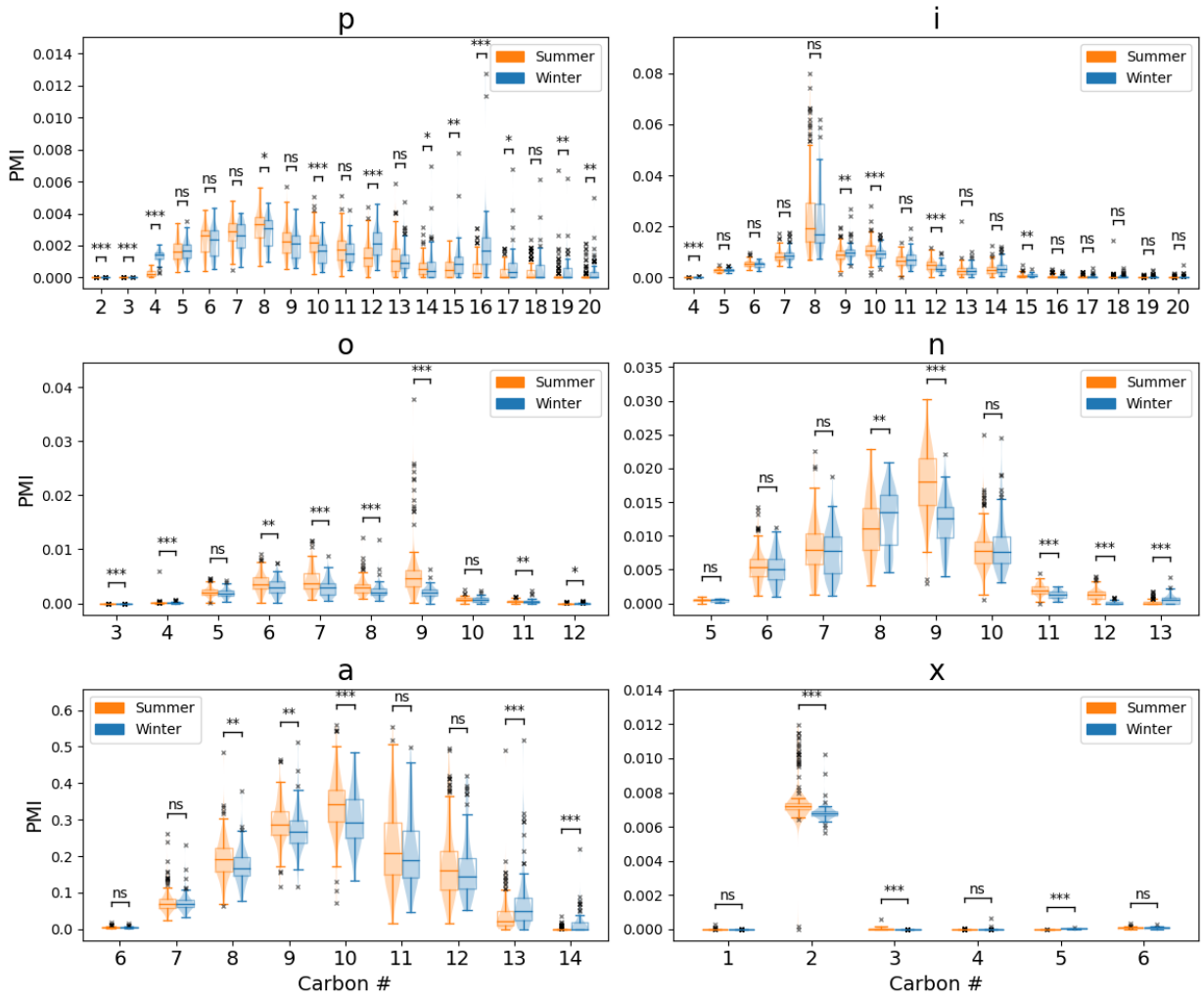

Figure 51. Comparison of PMI contributions of PIONAX groups between complete summer and winter datasets.

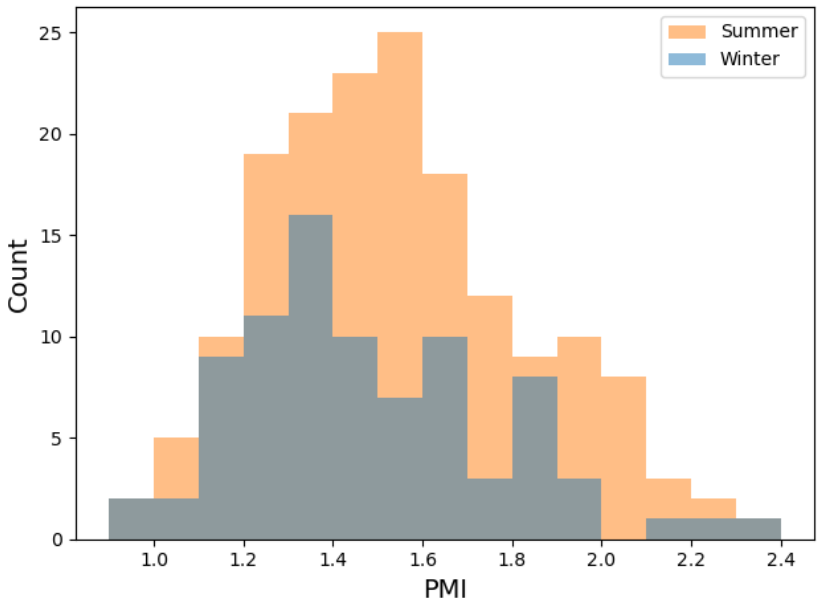

Figure 52. Histograms of fuel PMI for summer and winter datasets.

5.3.1. Comparison of distillation profiles between the two seasons

Figure 53 shows a comparison of distillation profiles between the two seasons as derived using different measurement methods. All methods show the same trend, namely that the winter fuels have more volatile content, with more material evaporating below 100ºC. This is especially pronounced for D86 and D7096 data. The composition-derived distillation profiles (D6730 and D8369) show most of the differences between the seasons below 75ºC. The differences in the amount of more volatile compounds correspond to the observed differences in the chemical composition, with the winter fuels having significantly higher content of lighter paraffins, such as butane (Section 5.3.1, Figure 47).

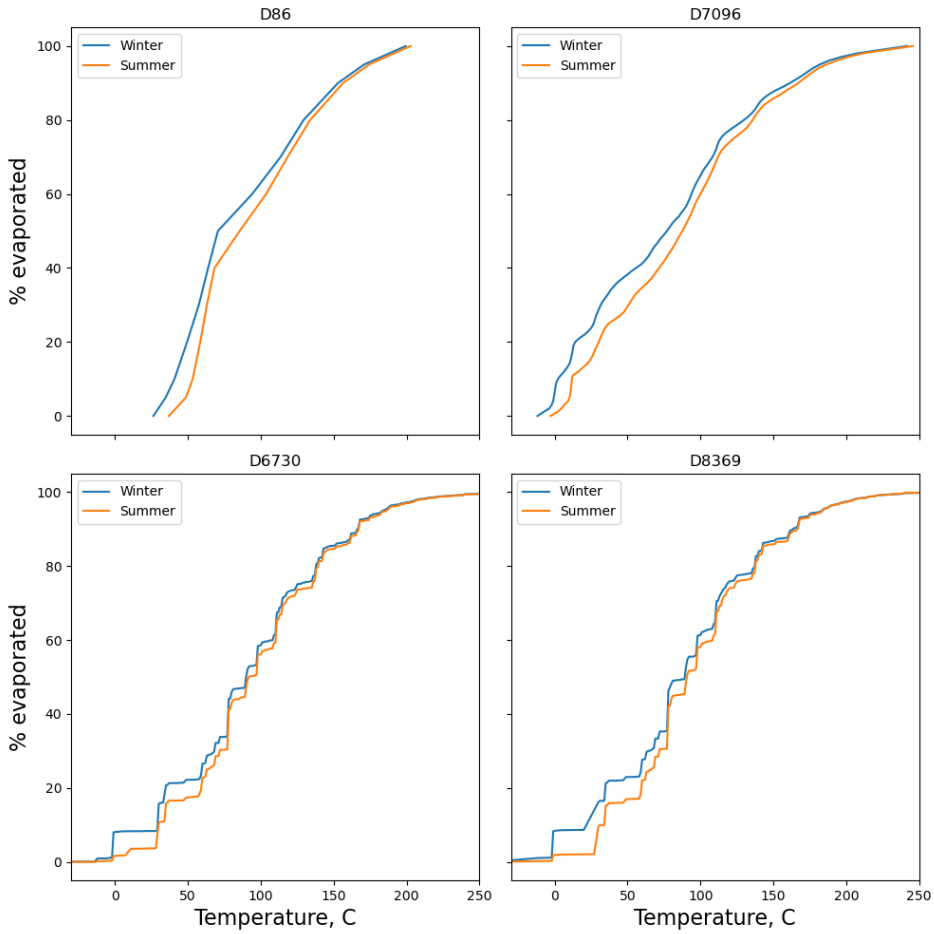

Figure 53. Comparison of average distillation profiles between summer and winter datasets for different measurement methods.

5.3.2. Comparison with the 2010 Honda study

The descriptive statistics of PMI values (both including and excluding the interpolated values) as well as those estimated for the US fuels of the Honda 2010 study are given in Table 7. The Honda study included 406 samples of US fuels collected in summer 2008 and summer 2009 and analyzed using the ASTM D6729 method. The statistics of the Honda study were obtained by interpolating the digitized cumulative data for that study and then calculating frequency distribution using the interpolated values. Thus, these values for the Honda study are only an approximation but do provide a means to assess how fuel PMI has changed since that study. Figure 54 show the cumulative PMI distributions for the summer and winter datasets as measured with the D6730, as well as those from the Honda 2010 study.

There is a clear and significant reduction in fuel PMIs since the Honda study. The most meaningful comparison with the Honda study is between PMI values that do not include the interpolated values, as it is unlikely that the Honda study used any interpolation for generic and unknown compounds. The mean PMI decreased from approximately 1.76 to 1.23-1.31. Since the mean PMI of the Honda study was estimated, it is probably safer to use the median point, which was 1.69 in the Honda study and 1.21-1.31 for D6730 for this study. Greater reductions are

observed for the higher end of the PMI distribution, as a tested by the "sharper" cumulative distributions of this study relative to that of the Honda study.

It should be noted that the chromatographic methods such as the ASTM D6730 have evolved significantly in recent years, both improving identification of compounds and extending the chromatographic range, all of which probably bias the current methods high relative to the methods used in the Honda study. The decrease in PMI, thus, is probably even larger than is reported here.

Table 7. Descriptive statistics of PMI values in winter and summer datasets of this study and those estimated for the US fuels of Honda 2010 study. "pmi": PMI values excluding interpolated values; 'pmi_int" PMI values including interpolated values.

	Winter				Summer				Honda 2010, US summer fuels
	D6730		D8369		D6730		D8369		
	pmi	pmi_int	pmi	pmi_int	pmi	pmi_int	pmi	pmi_int	
count	84	84	85	85	169	169	169	169	406
mean	1.23	1.48	0.96	1.15	1.31	1.54	1.05	1.27	1.76
std	0.20	0.28	0.16	0.23	0.22	0.29	0.19	0.27	0.46
min	0.81	0.91	0.61	0.70	0.70	0.89	0.41	0.51	0.80
25%	1.09	1.27	0.85	1.00	1.14	1.32	0.91	1.08	1.44
50%	1.21	1.42	0.96	1.11	1.31	1.52	1.07	1.26	1.69
75%	1.34	1.62	1.06	1.28	1.46	1.72	1.17	1.43	2.02
max	2.01	2.34	1.55	1.80	1.95	2.32	1.65	2.19	3.13

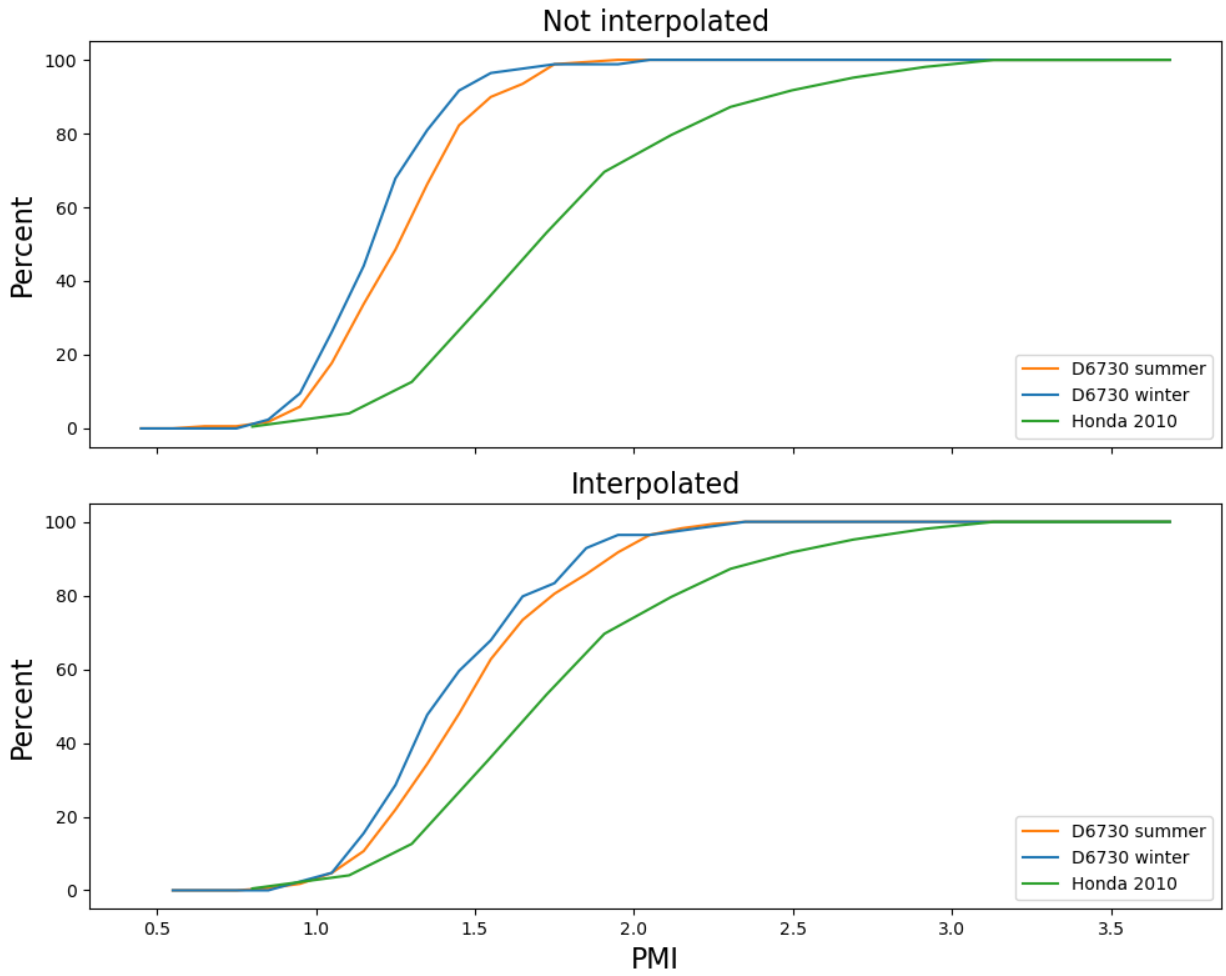

Figure 54. Cumulative distributions of PMI values for summer and winter datasets of this project and that of the Honda 2010 study.

6. Conclusions

Data from two fuel surveys (169 summer fuels from CRC project RW-121 and 85 winter fuels survey from CRC project RW-121-2) were analyzed to investigate differences in fuel composition, distillation profiles, and PMI among fuel grades, PADD regions, and seasons. Changes in fuel PMI relative to the Honda 2010 study were also investigated. The data included fuel chemical composition determined with ASTM D6730 and D8369, and distillation profiles measured using ASTM D7096 and ASTM D86. The data were harmonized to allow intercomparison between the methods.

While the D6730 and D8369 methods correlated well with each other in the total amounts of individual PIONAX groups, a notable bias between the two methods was observed, with

paraffins, isoparaffins, and oxygenates reported by the D6730 being lower than those reported by the D8369 method, while aromatics, olefins, and naphthenes showing the opposite trend. The reported per-compound weight contributions correlate well between the D6730 and D8369 methods for compounds with weight percent contributions above approximately 1%. The agreement between the two methods below 1 weight % is very poor, with individual values differing by up to two orders of magnitude. No evidence for coelution artifacts was observed. Overall, the differences between the methods are likely due to biases in calibration and the lower sensitivity of the VUV method.

The reason for the large discrepancy at small concentrations is not clear but could be a result of the baseline noise interference, differences in discriminating between partially co-eluting compounds, as well as compound misidentification by one or both methods.

A comparison of PMI values obtained using the D6730 and D8369 methods reveals a similar pattern, with a strong correlation between the two methods. However, the FID method consistently reports higher PMI values than the VUV method, with differences of 25% in summer and 28% in winter. The likely reason for the VUV method's underestimation of PMI relative to the FID method is its narrower RI range, which excludes heavier compounds that typically contribute significantly to the overall PMI.

Both D6730 and D8369 demonstrate remarkable proficiency in identifying compounds or assigning a compound class to a chromatographic peak. The VUV method identifies all compounds, while the FID method reports only 0.2 weight percent in summer and 0.1 weight percent in winter as unknown, with a negligible impact on the overall PMI. The contribution of generic compounds (those for which only the compound class was identified) is also minimal. According to D6730 data, generics accounted for only 1% of all aromatic weight. The highest contribution of generics was observed for naphthenes, where they constituted approximately 10%-13% (winter and summer, respectively) of the total weight of that compound group. Their contribution to the PMI was most significant for aromatics, amounting to about 5% of the aromatic-related PMI.

The distillation profiles showed variations among the four methods. The methods aligned better in the intermediate boiling range, with discrepancies increasing at both ends of the distillation profile. The composition-derived distillation profiles showed good agreement with each other, though with a minor bias throughout most of the range that became more pronounced at the low and high boiling point ranges. The SimDis method showed significant disagreement with the other two chromatography-based methods (D6730 and D8369), particularly in the lighter half of the distillation range, for reasons that remain unclear.

Chemical composition analysis revealed statistically significant differences across seasons, fuel grades, and PADD regions. Between fuel grades, the most notable differences (beyond the higher ethanol content in E15 grade) were the elevated levels of isooctane and toluene in premium unleaded gasoline compared to other grades. This aligns with the use of these compounds to achieve the higher (R+M)/2 rating required for premium grade gasoline. While compositional differences were observed between PADD regions, these variations were minimal and lacked statistical significance.

The primary seasonal difference in gasoline composition is the concentration of light compounds, particularly butane, which appears in higher quantities in winter formulations compared to summer ones. This variation aligns with Reid vapor pressure (RVP) regulations that impose lower limits in summer than winter. The seasonal difference is also evident in distillation profiles, which show summer gasoline to be less volatile than winter gasoline.

Aromatic compounds constitute the main contributor to fuel PMI, with C8-12 aromatics (those containing 8 to 12 carbon atoms) making the most significant contribution. In summer, premium unleaded fuels demonstrated statistically significantly lower PMI than regular and E15 grades. Regional analysis shows that fuels in PADD 2 (Midwest region) had statistically significantly higher PMI values compared to other regions, while differences among the other regions were not statistically significant. On average, summer fuels showed a slightly higher PMI (1.54) than winter fuels (1.48), though this difference lacks statistical significance at the 95% confidence level.

A notable and significant reduction in fuel PMIs has occurred since the Honda 2010 study. The median PMI value, which was 1.69 in the Honda study, decreased to 1.21 in winter and 1.31 in summer according to D6730 measurements, and to 0.85 and 0.91 respectively according to D8369. Mean values similarly declined from 1.76 to 1.23 (winter) and 1.31 (summer) for D6730, and to 0.96 (winter) and 1.05 (summer) for D8369. The reduction is even more pronounced at the higher end of the PMI distribution.

7. References

[1] Khalek, I., T. Bougher, and J. J. Jetter; Particle Emissions From a 2009 Gasoline Direct Injection Engine Using Different Commercially Available Fuels. SAE Technical Paper, 2010-01-2117, 2010.

[2] Stein, R. A., J. E. Anderson, and T. J. Wallington; An Overview of the Effects of Ethanol-Gasoline Blends on SI Engine Performance, Fuel Efficiency, and Emissions. SAE Int.J.Engines, 6, (1), 470-487. 2013.

[3] Chapman, E., Winston-Galant, M., Geng, P., Latigo, R., and A. Boehman; Alternative Fuel Property Correlations to the Honda Particulate Matter Index (PMI). SAE Technical Paper 2016-01-2250, 2016.

[4] Aikawa, K., Sakurai, T., and J.J. Jetter; Development of a Predictive Model for Gasoline Vehicle Particulate Matter Emissions. SAE Technical Paper 2010-01-2115, 2010.

[5] Chapman, E., Winston-Galant, M., Geng, P., and S. Pryor; Development of an Alternative Predictive Model for Gasoline Vehicle Particulate Matter and Particulate Number, SAE Technical Paper, 2019-01-1184 2019.

[6] Chapman, E., Winston-Galant, M., Geng, P., Latigo, R. et al., Alternative Fuel Property Correlations to the Honda Particulate Matter Index (PMI), SAE Technical Paper 2016-01-2250, 2016.

[7] Fatouraie, M., Frommherz, M., Mosburger, M., Chapman, E. et al.; Investigation of the Impact of Fuel Properties on Particulate Number Emission of a Modern Gasoline Direct Injection Engine, SAE Technical Paper, 2018-01-0358, 2018.

[8] An Improved Index for Particulate Matter Emissions (PME) Final Report for CRC Project RW-107-2. Coordinating Research Council, Alpharetta, GA. March 2021.

[9] ASTM International; Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100 Metre Capillary High Resolution Gas Chromatography. Designation: D6729 - 14, 1-51. 2016.

[10] ASTM International; Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100-Metre Capillary (with Precolumn) High-Resolution Gas Chromatography. Designation: D6730 - 22, Appendix X1. 2022.

[11] ASTM International; Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 50-Metre Capillary High Resolution Gas Chromatography1. Designation: D6733 - 01 (Reapproved 2011), 1-23. 2011.

[12] Enhanced Speciation of Gasoline, Final Report for CRC Project AVFL-29. Coordinating Research Council, Alpharetta, GA. June 2018.

[13] Geng, P., Collin, W., and Akers, V.; A New Predictive Vehicle Particulate Emissions Index Based on Gasoline Simulated Distillation, SAE Technical Paper 2022-01-0489, 2022

[14] Reilly, V., Goralski, S., Salyers, J., Geng, P. et al.; Correlation of Detailed Hydrocarbon Analysis with Simulated Distillation of US Market Gasoline Samples and its Effect on the SimDis Equation of Calculated Vehicle Particulate Emissions, SAE Technical Paper 2023-01-0298, 2023

[15] Effect of DHA Development on PMI Variability, Final Report for CRC Project AVFL-29-2. Coordinating Research Council, Alpharetta, GA. January 2022.

[16] Alliance for Automotive Innovation, North American Fuel Survey. Copyrighted 2022 and 2023.

[17] Validation of the New PM Index Formula: Phase 2, Appendix B. CRC Report No. RW-107-3a. Coordinating Research Council, Alpharetta, GA. September 2023. Accessed at https://crcao.org/wp-content/uploads/2023/10/PMECalculator.v4.0.02.xlsx