

EPA 600/R-05/032
February 2005

**Predicted Ground Water, Soil and Soil Gas Impacts
from US Gasolines, 2004
First Analysis of the Autumnal Data**

by

James W. Weaver
Ecosystems Research Division
National Exposure Research Laboratory
Athens, Georgia 30605

Lewis Jordan and Daniel B. Hall
Department of Statistics
University of Georgia
Athens, Georgia 30602

National Exposure Research Laboratory
Office of Research and Development
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

Notice

The U.S. Environmental Protection Agency through its Office of Research and Development funded and managed the research described here. It has been subjected to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Abstract

Ninety Six gasoline samples were collected from around the U.S. in Autumn 2004. The samples included regular and premium grade fuels, conventional and reformulated gasolines, high and low elevation samples, and fuel subject to state regulations (especially bans of methyl tert-butyl ether). A detailed hydrocarbon analysis was performed on each sample resulting in data set of approximately 300 chemicals per sample. Comparisons were made between several significant individual parameters. The results showed that benzene was, as required by the Clean Air Act, below 1 percent in areas that require reformulated gasoline (RFG). Higher benzene levels were found in some samples outside these areas. Methyl tert-butyl ether (MTBE) was found in use in RFG areas where there were no state bans in place. Where MTBE was banned, ethanol was found as the replacement oxygenate. Statistical analyses were performed on the entire suite of reported chemicals. These analyses were used to determine which gasolines were similar and if characteristics of the gasolines (conventional/reformulated, grade, elevation, and MTBE ban status) could be used to segregate the fuels into a decision tree. The statistical analyses showed that the gasolines were separable by these factors, but that differences due to state regulations or variability in the fuels make certain gasolines dissimilar from gasolines of similar characteristics. This feature was evident in the premium grade gasolines, presumably because refiners use different approaches to boost the fuel's octane number. Estimated solubilities and vapor pressures of the majority of the chemicals were used to predict the effective solubilities and gas phase concentrations of each chemical in a gasoline. Three types of low elevation, regular grade gasolines were included: conventional, MTBE-RFG and ethanol RFG. Pair-wise comparisons of these fuels showed predicted shifts in aqueous and gas phase concentrations.

Acknowledgments

The authors gratefully acknowledge the contributions of:

Hal White, US EPA, Office of Underground Storage Tanks;
Dr. Patricia Ellis, Delaware Department of Natural Resources and Environmental Conservation;
Dr. Michel Boufadel, Temple University, Civil Engineering;
Chris Englehart and Joe Haas, New York State Department of Environmental Conservation;
Dr. John T. Wilson, Eric Kleiner, and Fran Kremer, U.S. EPA, Office of Research and Development;
Jen Traub and Dr. Randall J. Charbeneau, The University of Texas at Austin, Department of Civil and Environmental Engineering;
Matt Jones, Levi Hein, Brad Thornhill, and Dr. Ken Rainwater, Texas Tech University, Department of Civil and Environmental Engineering;
Jeff Kuhn, Montana Department of Environmental Quality;
Chris Dufex and Mahesh Albuquerque, Colorado Division of Oil and Public Safety;
Andy Hess and Wally Moon, U.S. EPA Region 10, Seattle Washington;
Mark Restaino U.S. EPA Region 5, Chicago, Illinois;
Greg Lovato, Dr. Matthew Small, and Steve Linder, U.S. EPA Region 9, San Francisco, California;
Randall S. Weaver, Statham, Georgia;
Susan Swanson, The Cascade Group, Streetsboro, Ohio
Mike Birke and Robert Legg Southwest Research Institute, San Antonio, Texas
Drs. Fred Tillman, and Jay Choi, National Research Council, Athens, Georgia;
Dr. Drew Ekman, Dr. Mike Cyterski, Carlyn Haley, Pam Gunter, Dr. Said Hillal, Dr. Susan Richardson, Tim Rowan, David Spidle, and Dr Paul Winget, U.S. EPA Office of Research and Development, Athens, Georgia,
Sue Davis of U.S. EPA Office of Research and Development, Research Triangle Park, North Carolina, and
Rachel Claussen of U.S. EPA, Research Triangle Park, North Carolina.

Special thanks to Brandy Manders of U.S. EPA, Office of Research and Development, Athens, Georgia

Foreword

The National Exposure Research Laboratory's Ecosystems Research Division (ERD) in Athens, Georgia, conducts research on organic and inorganic chemicals, greenhouse gas biogeochemical cycles, and land use perturbations that create direct and indirect, chemical and non-chemical stresses, exposures, and potential risks to humans and ecosystems. ERD develops, tests, applies and provides technical support for exposure and ecosystem response models used for assessing and managing risks to humans and ecosystems, within a watershed / regional context.

The Regulatory Support Branch (RSB) conducts problem-driven and applied research, develops technology tools, and provides technical support to customer Program and Regional Offices, States, Municipalities, and Tribes. Models are distributed and supported via the EPA Center for Exposure Assessment Modeling (CEAM) and through access to Internet tools (www.epa.gov/athens/onsite).

ERD undertook a nationwide study of gasoline composition to generate insight and input data for risk assessment models, such as the Hydrocarbon Spill Screening Model (Weaver et al., 1994) and the OnSite on-line calculators (Weaver, 2004). Further, from the nationwide scope of the study, with its intent to include important geographic and regulatory zones, a set of relationships defining typical gasolines has been developed. This report describes results from the first analysis of data collected in the autumn of 2004. Future publications will present results from later sampling and more comprehensive analyses.

Rosemarie C. Russo, Ph.D.
Director
Ecosystems Research Division
Athens, Georgia

Leaking Underground Storage Tank Assessment Report Series

A series of research reports is planned to present data and models for leaking underground storage tank risk assessments. To date these include:

1. Gasoline Composition

Weaver, James W., Lewis Jordan and Daniel B. Hall, 2005, Predicted Ground Water, Soil and Soil Gas Impacts from US Gasolines, 2004: First Analysis of the Autumnal Data, United States Environmental Protection Agency, Washington, D.C., EPA/600/R-05/032.

2. Simulation Models

Weaver, James W., 2004, On-line Tools for Assessing Petroleum Releases, United States Environmental Protection Agency, Washington, D.C., EPA 600/R-04/101.

As more reports are added to the series, they may be found on EPA's web site at:
<http://www.epa.gov/athens/publications>.

Table of Contents

Notice	ii
Abstract	ii
Acknowledgments	iii
Foreword	iv
Leaking Underground Storage Tank Assessment Report Series	v
List of Figures	3
List of Tables	5
Abbreviations and Acronyms	7
Introduction	8
Literature and Purpose	9
Purpose	11
Approach and Methods	12
Results	18
Octane Number	18
Benzene	19
Methyl Tert-Butyl Ether	20
Tert-Butyl Alcohol	22
Tert-Amyl Methyl Ether	23
Oxygen Content	24
Ethanol	27
Brand	29
Similarities and Differences Among Gasolines	30
Cluster Analysis	30
Principal Components Analysis	41
Discriminant Analysis	49
Representative Gasolines	50
Cluster Analysis	52

Predicted Environmental Impacts of the Study Fuels on Water and Air	53
Predicted Aqueous Concentrations	54
Predicted Gas Phase Concentrations	63
Discussion and Conclusions	74
References	76
Appendix: Theoretical Basis of the Principal Components Analysis	79

List of Figures

Figure 1	Benzene data for RFG and conventional gasoline locales.	19
Figure 2	MTBE and benzene contents for all fuels.	20
Figure 3	MTBE content plotted against octane number.	21
Figure 4	Tert-Butyl Alcohol (TBA) content plotted against MTBE content, showing only small amounts of TBA present in these fuels.	22
Figure 5	Tert-Amyl Metyl Ether content plotted against octane number.	23
Figure 6	Composite weight of oxygen from all oxygenates and benzene concentrations.	24
Figure 7	Fuel oxygen content plotted against octane number for all samples.	26
Figure 8	Ethanol content plotted against octane number.	27
Figure 9	Ethanol content plotted against MTBE content showing the mutual exclusive nature of the usage of these chemicals.	27
Figure 10	Distribution of benzene by brand.	29
Figure 11	Dendrogram showing the cluster analysis' separation of 93 octane (left) from 87 (right) octane Georgia gasolines.	31
Figure 12	Dendrogram showing relationship between regular and premium gasolines from conventional gasoline areas in Georgia, Ohio and Oklahoma.	34
Figure 13	Dendrogram showing relationship between high elevation areas in Lubbock, Texas and Helena, Billings and Great Falls, Montana.	36
Figure 14	Dendrogram showing relationships between conventional gasolines from Georgia and RFG gasolines from Virginia.	36
Figure 15	Analysis of three RFG areas without MTBE bans.	37
Figure 16	Dendrogram of RFG gasolines from an MTBE-ban state (New York) and two states without bans.	38
Figure 17	Dendrogram of MTBE-ban states with State or Federal RFG requirements.	39
Figure 18	Dendrogram showing relationship between two MTBE ban, RFG gasoline areas.	41
Figure 19	Plot of the cumulative proportion of variation explained by PC number.	42
Figure 20	Scree graph of the variance estimate by PC number for the individual gasoline component data.	43
Figure 21	Principal Components plot of the 96 gasoline samples with respect to the first two principal components labeled by state.	46
Figure 22	Principal Components plot of the 96 gasoline samples with respect to the first two principal components labeled by octane number.	46
Figure 23	Plot of the 96 gasoline samples with respect to the first two principal components by gasoline brand.	48
Figure 24	Decision tree for gasolines typical of the data collected in this study. The parenthesis indicate where the cluster analysis indicated that samples from the various states are similar enough to group together.	50
Figure 25	Cluster analysis performed on the representative gasolines.	52
Figure 26	Effective solubility of hydrocarbon groups and oxygenates for the average conventional, low-elevation, no-MTBE-ban, regular gasolines.	55
Figure 27	Effective solubility of hydrocarbon groups and oxygenates for the average reformulated, low-elevation, no-MTBE-ban regular gasolines.	56

Figure 28 Effective Solubility for hydrocarbon groups and oxygenates for the Illinois regular gasolines (low elevation, MTBE ban, RFG). 59

Figure 29 Gas phase concentration of hydrocarbon groups and oxygenates for the average conventional, low-elevation, no-MTBE-ban regular gasolines. 63

Figure 30 Gas phase concentration of hydrocarbon groups and oxygenates for the average reformulated, low-elevation, no-MTBE-ban regular gasolines. 64

Figure 31 Gas phase concentration for hydrocarbon groups and oxygenates for the Illinois regular gasolines (low elevation, MTBE ban, RFG). 68

Figure 32 Scatter plot of predicted gas phase concentration versus the effective solubility of the average conventional, low-elevation, no-MTBE-ban, regular gasoline. 72

List of Tables

Table 1	Geographic locations, characteristics, and numbers of samples included in the study. Elevations less than 2500 ft are judged “low.” Fuel requirements	13
Table 2	ASTM D 6729-1 known pairs of co-eluting peaks.	15
Table 3	Percent by weight of unidentified compounds in analytical results.	16
Table 4	Specified chemicals included in the analysis.	16
Table 5	Hydrocarbon groups and numbers of member compounds.	16
Table 6	Octane ratings of sampled gasolines.	18
Table 7	Oxygen content of each oxygenated additive, based upon the number of carbon, hydrogen and oxygen atoms per mole of compound.	24
Table 8	Cluster centroids in weight % for Georgia 87 octane and 93 octane gasolines defined by all components present at 0.50 wt % or greater.	31
Table 9	Compounds that were present in <i>greater</i> abundance in Georgia premium than regular grades by at least 0.50 wt %.	33
Table 10	Compounds that were present in <i>lesser</i> abundance in Georgia premium gasolines than in regular grades by at least 0.50 wt %.	33
Table 11	Degree of similarity where all fuels of a given grade join in cluster analyses of individual locales (states).	35
Table 12	Statistics of the first eight PC from the PCA of the individual gasoline components including variance, proportion of the total variance estimate, and the cumulative proportion of variation explained.	42
Table 13	Gasoline components with the largest loadings in the first eight principal components including class, mean, standard deviation, and the rank of the standard deviation (from largest (1) to smallest(8)).	44
Table 14	Discriminant analysis results.	49
Table 15	Predicted <i>higher</i> aqueous constituent concentrations (>0.05 mg/L) of RFG (regular grade, no-MTBE-ban, low-elevation) in comparison to conventional gasoline.	57
Table 16	Predicted <i>lower</i> aqueous constituent concentrations (>0.05 mg/L) of RFG (regular grade, no MTBE ban, low elevation) in comparison to conventional gasoline.	59
Table 17	Predicted <i>higher</i> aqueous constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non-MTBE ban RFG gasoline.	60
Table 18	Predicted <i>lower</i> aqueous constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non-MTBE ban RFG gasoline.	61
Table 19	Predicted <i>higher</i> gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.	65
Table 20	(Continuation) Predicted <i>higher</i> gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.	67
Table 21	Predicted <i>lower</i> gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.	67
Table 22	Predicted <i>higher</i> gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.	69
Table 23	Predicted <i>lower</i> gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.	70

Table 24 (Continuation) Predicted *lower* gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.
..... 72

Abbreviations and Acronyms

BTEX	Benzene, Toluene, Ethylbenzene, Xylenes
CAAA	Clean Air Act Amendments
CG	Conventional gasoline
DHA	Detailed hydrocarbon analysis
DIPE	Di-Isopropyl Ether
DA	Discriminant Analysis
ETBE	Ethyl Tert-Butyl Ether
GC-MS	Gas chromatography/mass spectroscopy
HCA	Hierarchical Cluster Analysis
IUPAC	International Union of Pure and Applied Chemistry
LUST	Leaking underground storage tank
MTBE	Methyl Tert-Butyl Ether
M/Z	Mass to charge ratio
ON	Octane number
PC	Principal Component
PCA	Principal Components Analysis
RFG	Reformulated Gasoline
SPARC	SPARC Performs Automated Reasoning in Chemistry
TAE	Tert Amyl-Ethyl Ether
TAME	Tert Amyl-Methyl Ether
TAA	Tert Amyl Alcohol
TBA	Tert Butyl Alcohol
WO	Winter oxygenate

Introduction

Gasoline consists of numerous petroleum-derived chemicals and additives. At a fundamental level the composition determines the physical and operational properties of the fuel. Gasoline as a product, though, is defined in terms of these properties that include among others the octane number and the vapor pressure. The concentration of certain individual components remains important though as they may be regulated or mandated. The most notable of these are benzene and oxygenated additives. As a potential source of environmental contamination, however, the major consideration is the composition. Composition data are relatively scarce as historically the greatest concerns for chemical contamination of ground water at leaking underground storage tank (LUST) sites have been placed on benzene, methyl tert-butyl ether (MTBE) and other aromatics (toluene, ethylbenzene and the xylenes¹). The logic of this approach appears to follow from several factors: the status of benzene as a carcinogen, the low taste and odor threshold of MTBE, the relatively high concentrations in fuel of BTEX and, in some cases MTBE, and the solubilities and volatilities of these chemicals that allow them to form ground water plumes and/or vapor clouds in the subsurface.

A more fully multicomponent approach may be called for in certain circumstances. The non-MTBE oxygenates are becoming more prominent because of some states' bans on MTBE use, while the Federal oxygenate mandate remains in place. Ethyl Tert-Butyl Ether (ETBE), Tert Amyl-Methyl Ether (TAME), Tert Amyl Ethyl Ether (TAEE) and Di-Isopropyl Ether (DIPE) are possible ether oxygenates. Ethanol and Tert Butyl Alcohol (TBA) are possible alcohol oxygenates. As shifts are made away from MTBE as the dominant oxygenate, other changes in gasoline composition may occur. This might result in differing levels of various other contaminants in soil gas or ground water. Knowledge of shifts in composition can provide decision-makers with foreknowledge of potential impacts. A second reason for a multicomponent approach is that biodegradation is generally agreed upon to be electron-acceptor limited. Thus the pool of electron acceptors is available only for a finite mass of contaminants at most sites. The loading of all species contributes to usage of electron acceptors and may need to be considered. Multicomponent analysis may be useful for distinguishing among different gasolines (Stout et al., 2003) as part of an environmental forensics investigation. A fourth reason for studying multicomponent gasoline composition is that detailed characterization of gasolines may provide a means to predict the types of potential impacts that may occur. This need results from the very nature of leaking underground storage tanks: leaks occur unseen and undetected for years in many cases. It is clearly impossible to go back in time and measure the composition of the leaked fuel. Knowledge of composition is needed to model or estimate the risks associated with the release.

¹Benzene, toluene, ethylbenzene and xylene, taken together, are known as BTEX.

Literature and Purpose

The Clean Air Act Amendments (CAAA) of 1990 required the US EPA to develop oxygenated fuel regulations for cities that failed to meet carbon monoxide ambient air standards. Beginning in 1992 gasolines sold in these areas were required to contain a minimum of 2.7% oxygen by weight (Strickers, 2002). Reformulated gasolines (RFG) which contain oxygenates and meet other requirements were also mandated by the CAAA for certain ozone non-attainment areas after January 1, 1995. EPA regulations require that the benzene content of reformulated fuels be 1.0 % by volume or less when determined on a per gallon basis, or 0.95 % volume or less on average with a 1.3 % volume maximum per gallon (40 CFR 80, Subpart 41). The oxygen content must be 2.0 % by weight or more when determined on a per gallon basis, or 2.1 % weight or more on average with a 1.5 % by weight minimum per gallon (40 CFR 80 Subpart 41). A so-called anti-dumping provision mandates compliance in areas using conventional gasoline with baseline gasoline composition and properties based on 1990 gasoline production. State requirements, including MTBE bans are given in state laws for California (California Air Resources Board, 2003), Colorado (Colorado, 2002), and Washington (Washington State, 1996).

Data on gasoline composition have been collected by several groups. The Canadian Petroleum Products Institute (1994) performed a survey of Canadian gasolines during winter and summer of 1993. Their analysis included 44 compounds, which represented a fairly high fraction of most of the samples. Their study was divided by Canadian provinces, which gives the study an East/West geographical separation, but they did not report specific octane number nor elevation of the regular, mid and premium grade samples included in the 128 samples per season. Only two of the samples contained a significant amount of MTBE.

Other, less intensive, sources of data include the generalized composition reported by Gustafson et al. (1997), and Environment Canada (Jokuty et al., 1999), which contains six examples of detailed chemical composition of gasoline among 450 petroleum products and crude oils in total. Some states perform and publish surveys of gasoline composition which usually focus on benzene, BTEX and oxygenates (e.g., Maine Department of Environmental Protection, 2005)

U.S. EPA collects certain data on reformulated gasoline including the amount of benzene and oxygenates (US EPA, 2005). EPA collates data generated by industry collected in summer and winter surveys. Summaries of the data are available on the EPA web site for the years 1995 to 2003. The surveys may contain as many as 10,000 gasoline retail samples that are collected during one-week survey periods by the RFG Survey Association (U.S. EPA, 2005). Prior to 1998, the surveys reported only total oxygen and oxygenate content, benzene content, aromatics content and Reid vapor pressure. In 1998 and subsequent years, the surveys also reported sulfur, olefins and certain distillation properties. Separately EPA compiles compliance data from producers and importers. Average values of parameters are available for summer and winter composition for the years 1997-2002 for both conventional and reformulated gasolines as national averages.

A voluntary industry consortium produces semiannual Petroleum Product Survey reports and is based in Bartlesville, Oklahoma. This group collects data on benzene, oxygenates, the boiling point distribution, vapor pressure, octane number and other properties in annual winter and summer surveys of U.S. gasolines (e.g., Dickson 2004a, 2004b). These surveys have been conducted since the 1930s and each may contain data from as many as 800 to 900 samples. Winter and summer sampling periods last from December through February for winter samples and June through August for the summer samples.

The octane number (ON) of a fuel is an important measure of its quality (Strickers, 2002). The typical octane number of gasolines produced in the 1920s was 40-60. It has increased since in order to meet performance needs of modern engines, namely to allow higher compression ratios without pre-ignition of fuel and the resultant engine knock. The octane number is determined by measuring the combustion efficiency of a fuel in comparison to a prescribed mixture of isooctane (2,2,4-trimethylpentane) and n-heptane. Of these, the first has an ON of 100, and the second of 0, because of their ability or lack of ability, respectively, to prevent fuel pre-ignition. Thus the relative combination of these represents chemicals either suppressing or enhancing engine knock. The octane number posted at gasoline stations in the U.S. is an average of the Research Octane Number and the Motor Octane Number, although many other variations on Octane Number exist and are used for various purposes (Chevron U.S.A., Inc, 2005). Higher ON is provided to fuel by increased amounts of aromatics, olefins and iso-paraffins (Strickers, 2002)². In general, branched hydrocarbons and aromatics are more knock resistant than unbranched paraffins (ASTM, 1958, Meusinger and Moros, 2001). This characteristic is illustrated in the archetypical octane enhancer 2,2,4-trimethylpentane or isooctane.

The environmental impacts of fuels depend, in part, on their chemical properties. Data on chemical properties are available from a variety of sources. There are no “complete” data sets that include all hydrocarbons of interest. Further, the majority of property data are collected at temperatures of 25 °C and data for shallow subsurface temperatures (5 °C to 25 °C) are unavailable. The International Union of Pure and Applied Chemistry (IUPAC) published two series of hydrocarbon solubility data (Shaw, 1989a, 1989b) that include hydrocarbons from C₅ to C₇ and C₈ to C₃₆, respectively. For some well-studied compounds such as benzene, there were multiple sets of solubilities reported over a range of temperatures. Depending upon the amount and quality of the available data, IUPAC evaluated the data and provided recommended, best or reported values.

Gustafson et al. (1997) developed parameter correlations for solubility, vapor pressure and other properties based upon the equivalent carbon number of the compound. The correlations were developed for aromatic and aliphatic constituents of fuels. Use of the correlations requires estimation of the equivalent carbon number via the boiling point of the

²The oil industry uses the terms paraffins to refer to what are otherwise known as alkanes, olefins for alkenes and alkynes, and naphthenes for cycloalkanes.

chemical.

Because recommended values are not available for all desired properties of all chemicals at all temperatures of interest, estimated properties were used instead. The SPARC Performs Automated Reasoning in Chemistry (SPARC³) property estimator (Karikhoff et al., 1991, Hillal et al., 2003) uses artificial intelligence, molecular structure and large training datasets (8000 data points) to generate property estimates.

Purpose

The purpose of this study was to generate data on gasoline composition that could be used as inputs to risk assessment models for ground water or indoor air contamination. The approach taken was to characterize the chemical composition of gasolines from various locations around the U.S, using as detailed a method as practical. Unlike the established EPA/Industry consortia, the purpose here was to collect as much data as possible on the individual chemicals composing the fuels. This approach was chosen because environmental impacts from fuels depend on the specifics of their composition and the desire was to evaluate the impacts from the greatest number of hydrocarbons as possible.

³SPARC is available at <http://ibmlc2.chem.uga.edu/sparc>.

Approach and Methods

Gasolines were collected by sampling directly from retail pumps using the following general protocol. A list of specified brands was used to choose two to four gas stations within a 1 mile to 2 mile radius of each other. Select brands were used because of the possible, but unlikely, prospect that the petroleum-derived fractions of gasoline were brand specific. The chosen stations were to be “clean” in appearance to minimize the potential for inclusion of off-spec product in the study. Multiple stations were selected to provide close geographic clustering and, where possible, stations across the street from each other were included. These were included specifically to test for similarity among different brand names. To avoid mixing of product in delivery lines, the stations were required to have three separate hoses for dispensing gasoline.

The geographic locations that were chosen for the study were intended to include a mixture of locales using conventional or reformulated gasoline (RFG), locales where ethanol usage was likely, states with MTBE bans, low and high elevations, rural and urban areas, and a balance of East and West coast samples (Table 1). The designation of RFG areas was based on information from EPA that was updated as of February 23, 2004 (US EPA, 2005) and applies to the specific town or city included (as Federal RFG requirements are not state-wide). A state winter oxygenate requirement was in effect in Colorado (Colorado, 2002). Other requirements for gasoline include limitations on Reid vapor pressure and sulfur content in some areas of the country. Since these two were not included in the measured data, there is no further discussion. The varying requirements may, however, cause changes in other components of gasoline that are included in this study. Information on MTBE bans was provided by the American Petroleum Institute (API, 2004).

Summer and winter samples were to be taken. The number of samples was limited by the available project funding that meant roughly 120 samples in both winter and summer sampling seasons would be analyzed. Because of several reasons, sampling was delayed and the data presented in this report were collected between September and November 2004 and thus likely contain both winter and summer samples.⁴ This “autumnal” sample contains 96 samples, most of which are likely to be winter samples.

Regular and Premium gasoline were dispensed into separate 1 gallon gasoline cans and transferred to 20 ml scintillation vials. The vials were filled to the bottom of the threading: (1) to prevent gasoline from flowing over the outside of the vial, (2) to prevent breakage of vials due to thermal expansion, and (3) to comply with one of the requirements of the small quantity exceptions to dangerous good shipping under 49 CFR 173.4. The vials were then overnight shipped in chilled containers for storage in Athens, Georgia. In Athens, samples were then selected for analysis and shipped to The Cascade Group (TCG) for analysis. All analyses were

⁴Planned future sampling will provide true summer and winter samples and the data reported here will be combined with the appropriate summer or winter data.

completed by TCG within two weeks of receiving the samples.

The samples were analyzed by ASTM D 6729-1, a high resolution, gas chromatograph method. The method uses a 100 meter capillary column and flame ionization detector. This method was developed for the determination of individual hydrocarbon components of spark-ignition engine fuels, including alcohol and ether oxygenated additives, all with boiling points below 225 °C. Since a typical gasoline is a mixture of over 400 components, it would be impractical if not impossible to impose data quality indicators on each analyte of interest. Therefore, one component from each of the functional groups was tracked to assess the overall quality of the analytical performance. The laboratory routinely monitored the repeatability and reproducibility of its analysis. The repeatability was monitored through the use of laboratory replicates at the rate of one per 10 samples. Reproducibility was monitored through the use of a quality control sample analyzed at the rate at least one per 15 samples. Potential matrix effects were monitored by the use of a spiked sample. Every 20th sample was spiked with 1% each of the following components: Ethylbenzene, MTBE and 2,2,4-Trimethylpentane. The quality control samples were plotted on individual control charts and the upper and lower control limits were determined in accordance with laboratory protocols.

Using this method to identify gasoline components relies on retention time. As a result, there is a possibility of co-elution of some peaks. Although a majority of the hydrocarbons can be clearly resolved, there is a possibility of co-elution of compounds with olefins above C₇. Table 2 lists the potentially co-eluting compounds that are identified in ASTM D 6729-1. The method notes that the list is not exhaustive. Although co-elution raises a potential problem in interpreting the results, two factors mitigate this problem. First is that ASTM D 6729-1, Table 4, indicates that one compound is predominant. Of the compounds listed in Table 4, most are not of specific interest, with the exception of toluene. The second mitigating factor is that the co-eluting compounds listed in Table A1.1 of ASTM D 6729-1, are *usually* present at no more than 1000 ppm (0.1% by mass). Thus reported concentrations of MTBE and ETBE that themselves are above 0.1 represent likely masses of these chemicals and not the co-eluting chemicals. A compound that co-eluted with ETBE, Methylcyclopentane, was found to interfere with the ETBE results from ASTM D 6829-1. Despite any mitigating factors, the nature of the ETBE results, in particular, requires that confirmation by other means be performed. Confirmation is planned for future samples and thus the ETBE data were omitted from this report.

Table 1 Geographic locations, characteristics, and numbers of samples included in the study. Elevations less than 2500 ft are judged “low.” Fuel requirements for reformulated gasoline (RFG), winter oxygenates (WO), and conventional gasoline (CG) are provided on a location by location basis.

State	Locale(s)	Elevation	Fuel Requirements	MTBE ban	No. of Samples	
					Regular	Premium
California	Los Angeles (Huntington Beach, Irvine)	Low	RFG	Yes	5	3
Colorado	Denver	High	WO	Yes	3	3
Georgia	Athens (Bogart, Watkinsville)	Low	CG	No	4	5
Delaware	Dover	Low	RFG	No	3	3
Illinois	Chicago (Batavia)	Low	RFG	Yes	3	3
Montana	Billings, Great Falls, Helena	High	CG	No	6	6
New York	Long Island (Centereach, Merrick, Riverhead)	Low	RFG	Yes	6	6
Ohio	Cincinnati	Low	CG	No	3	3
Oklahoma	Ada	Low	CG	No	3	3
Pennsylvania/New Jersey	Philadelphia (Bala Cynwyd, PA, Pennschawken, NJ)	Low	RFG	No	2	3
Texas	High Plains (Lubbock, Wolfforth)	High	CG	No	3	3
Virginia	Arlington	Low	RFG	Yes	4	4
Washington	Seattle (Port Orchard)	Low	GG	Yes	3	3
<i>Total</i>					48	48

Table 2 ASTM D 6729-1 known pairs of co-eluting peaks.

Predominant compound	Coeluting Compound(s)	
	Chemical	Notes on Co-eluting compound
<i>Co-eluting pairs listed in Table 4 of D 6729</i>		
3,3-dimethylpentane	5-methyl-1-hexane	
2-methylhexane	C ₇ -olefin	
2,5-dimethylhexane	C ₈ -olefin	
3,3-dimethylhexane	C ₈ -olefin	
toluene	2,3,3-trimethylpentane	if 2,3,3-trimethylpentane > 5 times the toluene concentration
1,1,2-trimethylcyclopentane	C ₇ -triolefin	
C ₈ -diolefin	C ₈ -paraffin	
4-methyloctane	C ₉ -olefin	
1,2,3,4-tetramethylbenzene	C ₁₁ -aromatic	
<i>Co-eluting pairs listed only in notes to table A1.1 of D 6729</i>		
n-propanol	3-methyl-1-pentene	usually < 1000 ppm
MTBE	2,3-Dimethyl-1-butene	usually < 1000 ppm
MSBE (Methyl sec-butyl ether)	1-hexene	usually < 1000 ppm
ETBE	2,3-Dimethyl-1,3-butadiene	usually < 1000 ppm
isobutanol	4,4-Dimethyl-1-pentene	usually < 1000 ppm

The results included identification of 312 chemicals in gasoline, of these 27 were identified to only an isomer (i.e., C₁₁-Aromatic, C₁₀-Iso-Paraffin, etc.). Unknown chemicals were grouped into five categories: unidentified, unidentified-aromatic, unidentified-Iso-Paraffin, unidentified-naphthenes, and unidentified-olefins. Table 3 identifies the minimum, average and maximum amounts of each of these categories of unknowns. Of the remaining chemicals in gasoline, two sets of chemicals were specifically included in the analysis: the first group included BTEX, trimethylbenzenes and oxygenates, the second included the 44 chemicals used

in the Canadian Petroleum Products Institute study of 1993 (Table 4). The remainder of chemicals were selected from the analyte list of ASTM D 6729-1. In addition the chemicals were classified by hydrocarbon group (Table 5).

Table 3 Percent by weight of unidentified compounds in analytical results.

Group	Minimum	Average	Maximum
Unidentified	1.87	3.933	9.987
Unidentified-aromatics	0.004	0.049	0.094
Unidentified-ISO-Paraffins	0.088	0.333	0.584
Unidentified-naphthenes	0.08	0.267	0.801
Unidentified-olefins	0.021	0.082	0.202

Table 4 Specified chemicals included in the analysis.

Aromatics and Oxygenates	Additional CPPI (1994) Study Chemicals	
Benzene	1,3 Butadiene	Cyclopentane
Toluene	iso-Butane	Methylcyclopentane
Ethylbenzene	iso-Pentane	Cyclohexane
o-, m-, p-xylene	2,2-Dimethylbutane	Methylcyclohexane
1,3,5-Trimethylbenzene	2,3-Dimethylbutane	Naphthalene
1,2,4-Trimethylbenzene	2-Methylpentane	2-Methylnaphthalene
Methanol	3-Methylpentane	1-Methylnaphthalene
Ethanol	2,4-Dimethylpentane	trans-Butene-2
t-Butanol	2-Methylhexane	cis-Butene-2
Methyl tert-butyl ether	3-Methylhexane	trans-Pentene-2
Di-isopropyl ether	2,2,4-Trimethylpentane	cis-Pentene-2
Ethyl-tert butyl ether	2,4-Dimethylhexane	n-Butane
Tert-amyl alcohol	2,3,4-Trimethylpentane	n-Pentane
Tert-amyl methyl ether	2,3,3-Trimethylpentane	n-Hexane
Tert-amyl ethyl ether	2,3-Dimethylhexane	n-Heptane
	3-Methylheptane	
	2-Methylbutene-1	
	2-Methylbutene-2	
	1-Methyl-3-ethylbenzene	
	1-Methyl-4-ethylbenzene	
	1-Methyl-2-ethylbenzene	

Table 5 Hydrocarbon groups and numbers of member compounds.

Group	Members	Group	Members
Di/Bicyclo Naphthenes	1	Naphthenes	4
Di-Olefins	19	Napheheno/Olefino-Benzenes	3
Indenes	1	Naphtheno-Olefins	4
Iso-Paraffins	76	n-Olefins	37
Iso-Olefins	53	Oxygenates	12
Mono-Aromatics	48	Paraffins	12

SPARC was used to generate estimates of the boiling points, and, for temperatures from 0°C to 25°C in 5 °C increments, the vapor pressures, infinite dilution activity coefficients, and solubilities of the chemicals reported in the gasoline samples. Of the 285 chemicals identified in the analysis, SPARC was able to determine the properties for 281. These values were used in the assessment of effective solubility and gas phase concentration that follows.

Results

The data were evaluated in two ways. The first method was through comparisons of concentration of single components against the octane number or another component. In almost all cases scatter plots were used to get a sense of the abundance of various components under various circumstances. The second set of analyses were performed using the entire suite of chemicals reported for each sample. Because of the large number of gasoline components, statistical methods were used: hierarchical cluster analysis (HCA), discriminant analysis (DA), and principal components analysis (PCA). The results from these analyses provided assessment of the variation among the gasolines and formed the basis for the decision tree presented below.

Octane Number

Data collection was designed to obtain samples of regular and premium gasolines. The octane ratings of these fuels vary, largely because of elevation. Lower octane fuels are sold at higher elevations. Because of larger numbers of low elevation samples, the most common regular and premium gasoline octane numbers were 87 and 93, respectively (Table 6).

Table 6 Octane ratings of sampled gasolines.

Grade	Octane Number	Conventional	Reformulated	Total
Regular	85.0	0	3	3
	85.5	4	0	4
	86	2	0	2
	87	15	22	37
	87.5	1	0	1
	89	0	1	1
Premium	90	2	0	2
	91	9	5	14
	92	3	1	4
	93	9	19	28
Totals		45	51	96

Benzene

The benzene data showed variation at all octane levels (Figure 1). Because of mandates of the Clean Air Act (CAA), benzene levels in RFG areas were approximately 1% or less⁵. In conventional gasolines benzene varied at levels that were generally between 0.5% and 3% by weight. The higher levels (greater than 2%) were found over the whole octane range (85.5 to 93). The triangles indicating winter oxygenate and high benzene were from the Colorado samples, where a wintertime oxygenate requirement is imposed (Colorado, 2002).

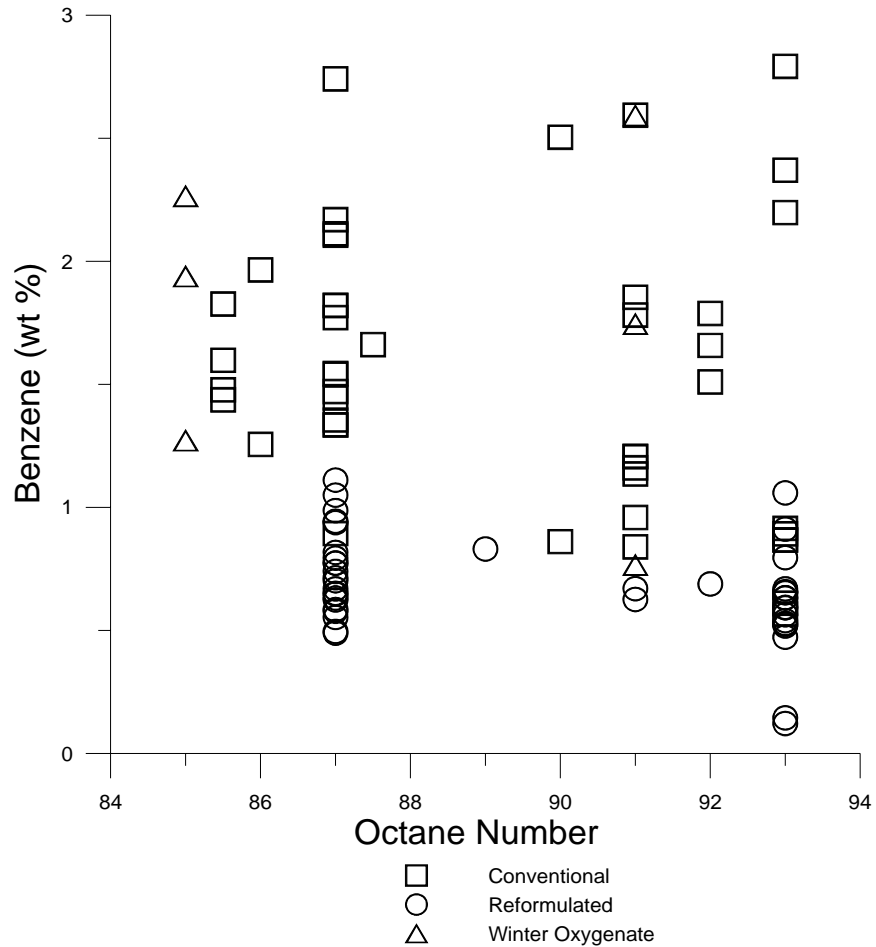


Figure 1 Benzene data for RFG and conventional gasoline locales.

⁵The federal FRG mandate on benzene is given in volume percent, as opposed to the weight percent used for these data.

Methyl Tert-Butyl Ether

The MTBE concentrations plotted against the benzene concentrations are shown in Figure 2. The fuels with high MTBE concentrations all occur with benzene concentrations less than 1%. This follows from requirements of the RFG program where benzene concentrations must be less than one percent by volume. Five conventional gasoline samples had MTBE concentrations of 2.8% to 4.2%. These concentrations could be the result of MTBE use as an octane enhancer, mixing of fuel types, or other reasons. All of these samples, however, were premium gasolines obtained in Georgia, suggesting the deliberate use of the oxygenate as an octane enhancer. At benzene concentrations above 1%, all the MTBE concentrations are low and most of the samples come from conventional gasoline areas.

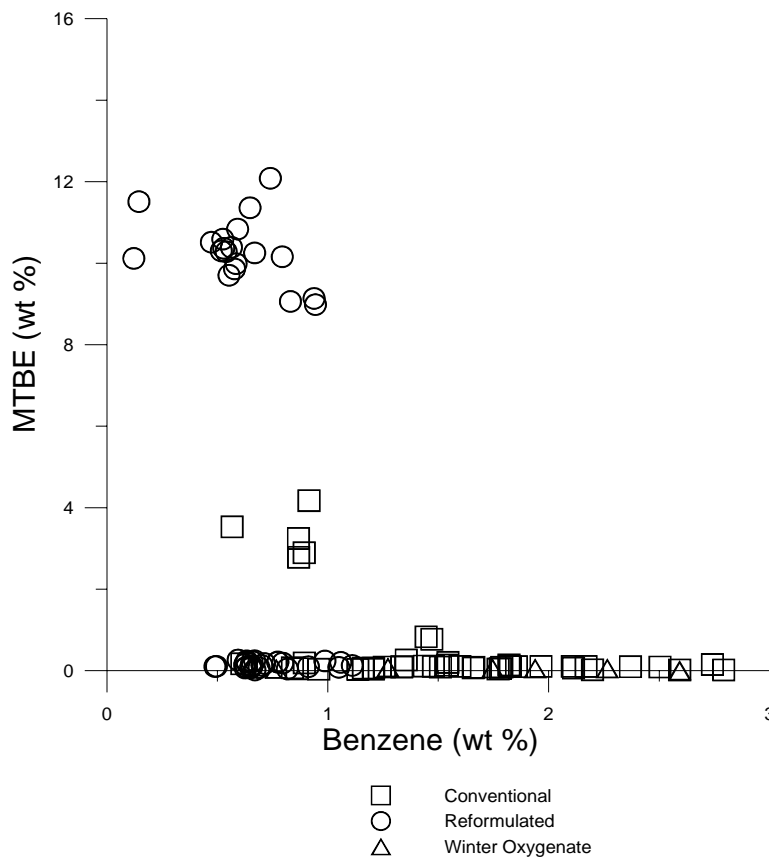


Figure 2 MTBE and benzene contents for all fuels.

When compared with the octane number (Figure 3), the MTBE concentration appeared to depend on RFG requirements more than octane number. MTBE is high where RFG requirements are in place and there is no MTBE ban. Otherwise the MTBE content is fairly low, with the exception of the Georgia premium samples.

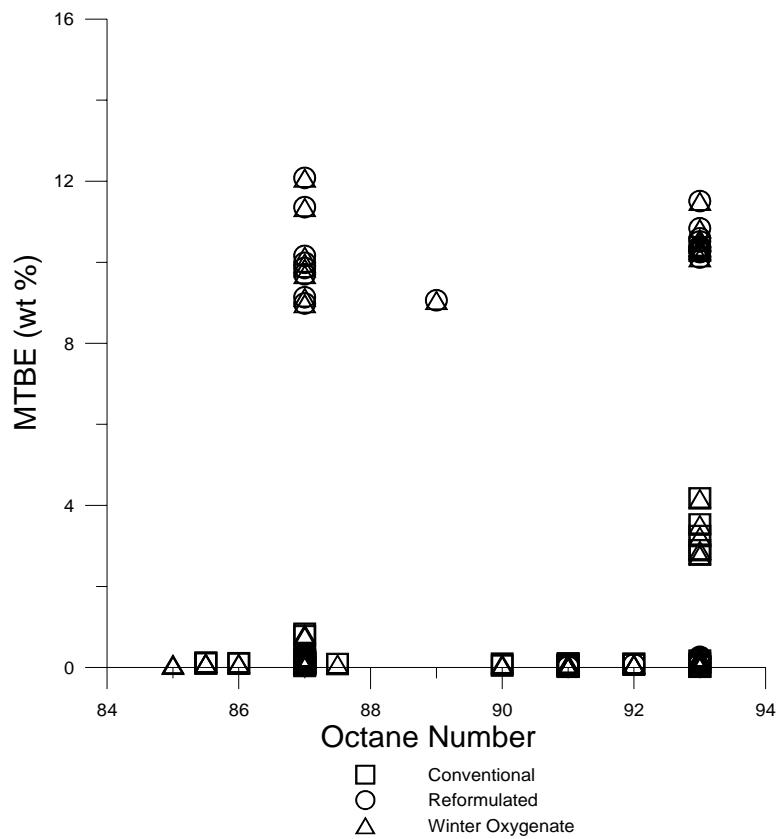


Figure 3 MTBE content plotted against octane number.

Tert-Butyl Alcohol

Figure 4 shows that the Tert-Butyl Alcohol (TBA) concentrations plotted against MTBE content of the fuels. All TBA concentrations were about 0.1 % or less. The higher values occurred with higher MTBE content, indicating some tendency for TBA to increase with MTBE.

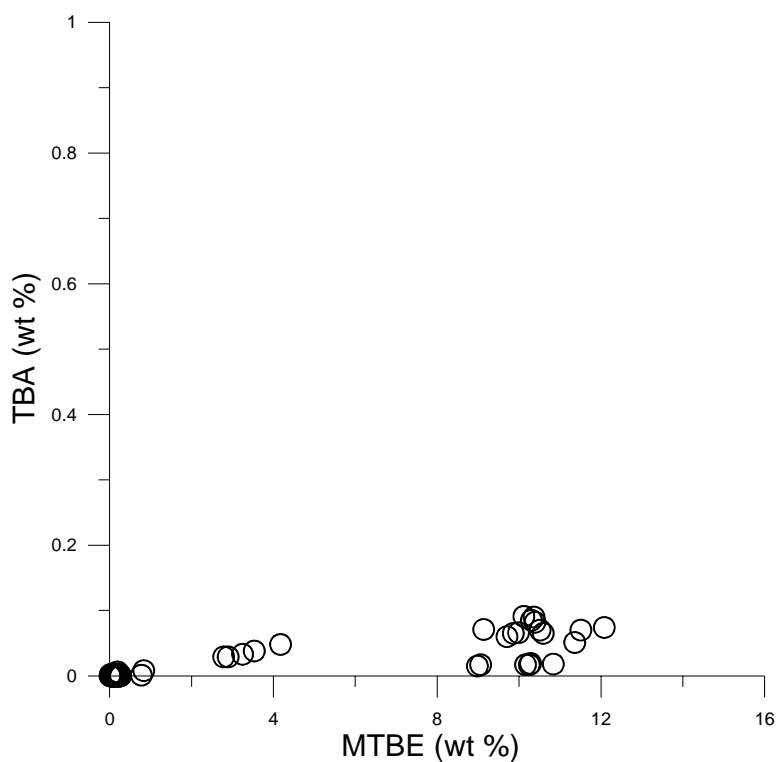


Figure 4 Tert-Butyl Alcohol (TBA) content plotted against MTBE content, showing only small amounts of TBA present in these fuels.

Tert-Amyl Methyl Ether

Tert-Methyl Amyl Ether (TAME) was found at only low concentrations (< 1%) in the fuel samples (Figure 5).

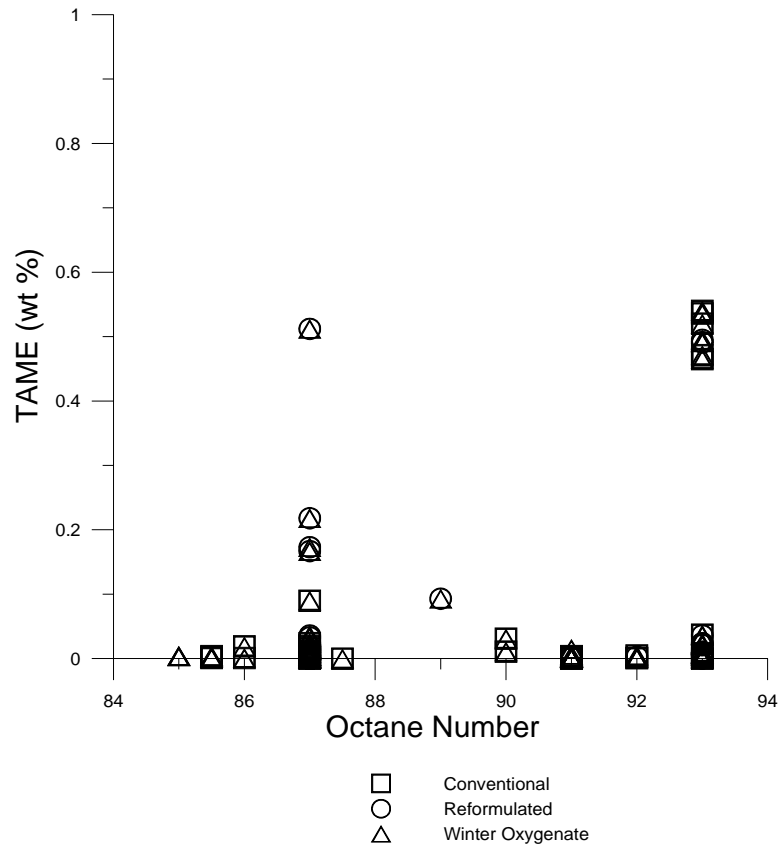


Figure 5 Tert-Amyl Methyl Ether content plotted against octane number.

Oxygen Content

The oxygen content of the fuels was supplied by several compounds including MTBE, ethanol and ETBE. Small amounts of TAME and DIPE were found in some samples. The total oxygen content of each sample was calculated by summing the contributions from each oxygenate. Table 7 lists the fraction of oxygen supplied by each. Because the weight fraction of oxygen in each additive varies, the amount of the compound needed to meet the RFG requirements also varies. Because the ETBE data were in doubt, ETBE was omitted from the results presented in Figure 6. The actual oxygen content of the fuels may be higher than indicated because this omission.

Table 7 Oxygen content of each oxygenated additive, based upon the number of carbon, hydrogen and oxygen atoms per mole of compound.

Oxygenate	Number of Atoms			Weight Fraction of Oxygen
	Carbon 12 g/mole	Hydrogen 1 g/mole	Oxygen 16 g/mole	
Methanol	1	4	1	0.50
Ethanol	2	6	1	0.35
Tert-Butyl Alcohol	4	10	1	0.22
Methyl Tert-Butyl Ether	5	12	1	0.18
Di-Isopropyl ether	6	14	1	0.16
Ethyl Tert-Butyl Ether	6	14	1	0.16
Tert-Amyl Alcohol	5	12	1	0.18
Tert-Amyl Methyl Ether	6	14	1	0.16
Tert-Amyl Ethyl Ether	7	16	1	0.14

The oxygen content of these additive were summed and plotted against the benzene concentration in Figure 6. RFG fuels show oxygen contents above 2 %. The majority of these had benzene concentrations less than one. Five samples had total oxygen about 3% and had high benzene concentration (above 1%). These were collected in Denver where state requirements mandate the oxygen content at this level. Conventional gasoline samples showed oxygen content below 1%, but not zero.

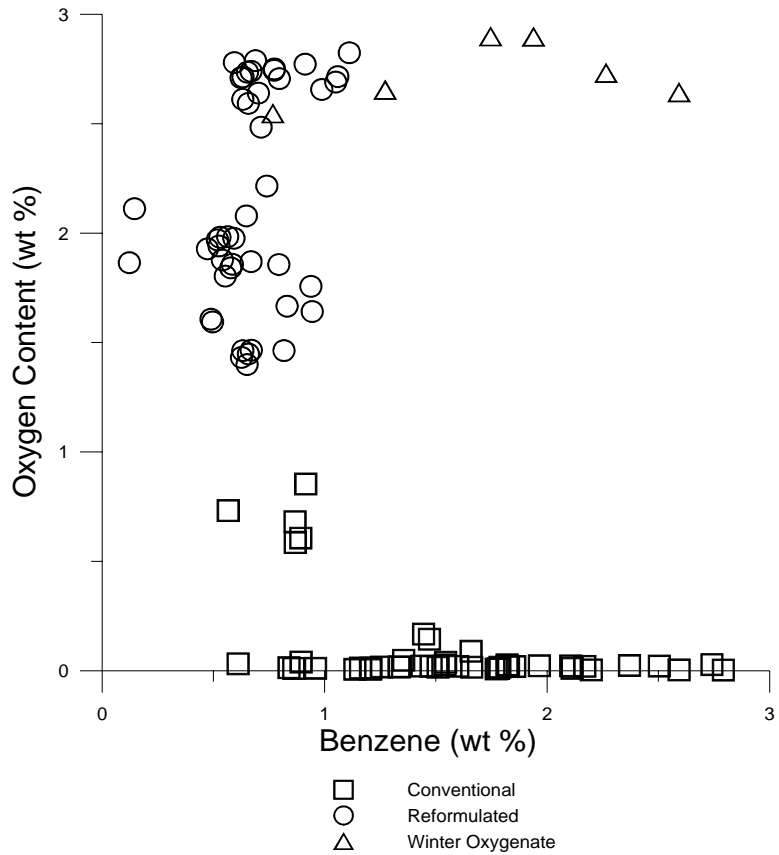


Figure 6 Composite weight of oxygen from all oxygenates and benzene concentrations.

Figure 7 shows the oxygen content (wt %) plotted against the octane number for both reformulated and conventional gasoline locations. This plot also shows that oxygen content is high in RFG areas and that lesser amounts ($< 1\%$) were found in conventional gasolines.

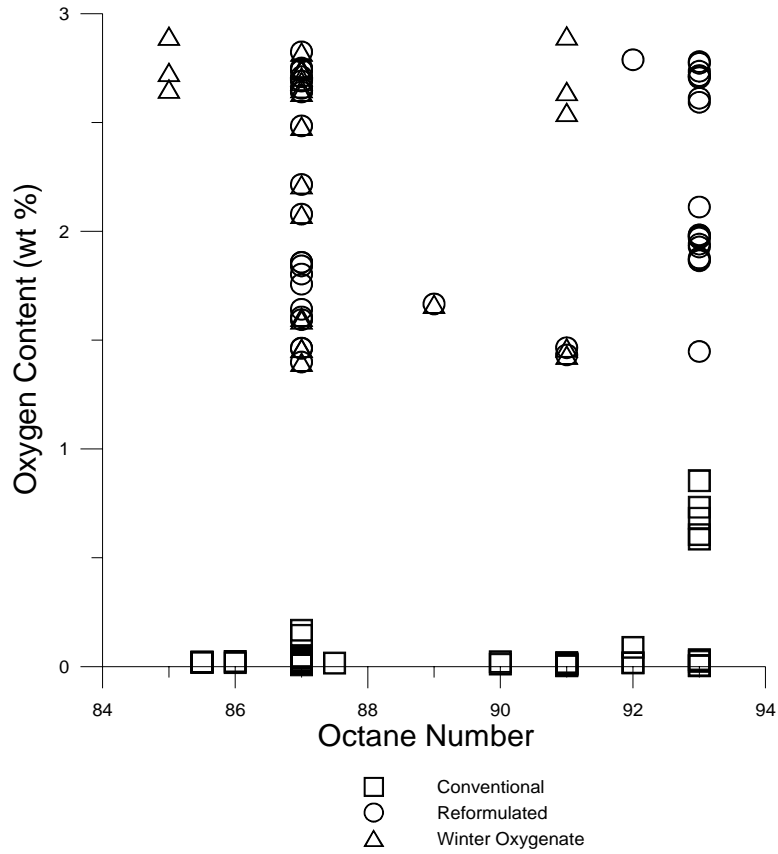


Figure 7 Fuel oxygen content plotted against octane number for all samples.

Ethanol

Ethanol has replaced MTBE as the principal oxygenate in three states included in the Study: New York, Illinois, and California. Figure 8 shows the ethanol content in the fuels. The ethanol content was high only in RFG locales, indicating little incidental presence where an oxygenate was not required. Figure 9 shows that MTBE and ethanol use were mutually exclusive, as the presence of one was associated with the absence of the other.

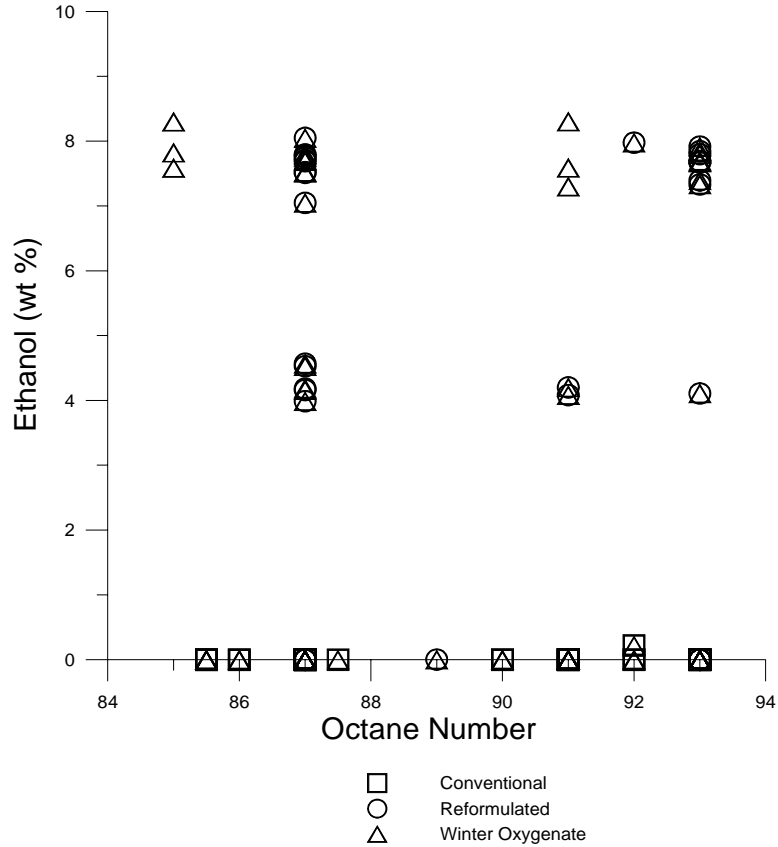


Figure 8 Ethanol content plotted against octane number.

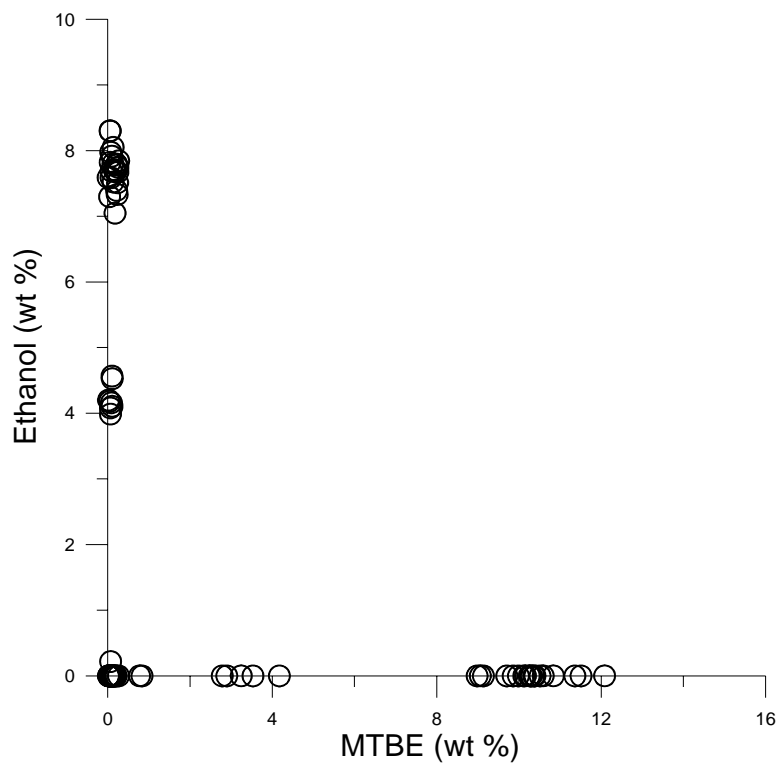


Figure 9 Ethanol content plotted against MTBE content showing the mutual exclusive nature of the usage of these chemicals.

Brand

The gasolines sampled were sold under seventeen brand names. There is little evident distinction among brands as indicated in Figure 10. This plot indicates that the benzene content varies over roughly the same range regardless of brand. Other indicators of the lack of significance of brand were generated by the cluster analysis and the principal components analysis presented later.

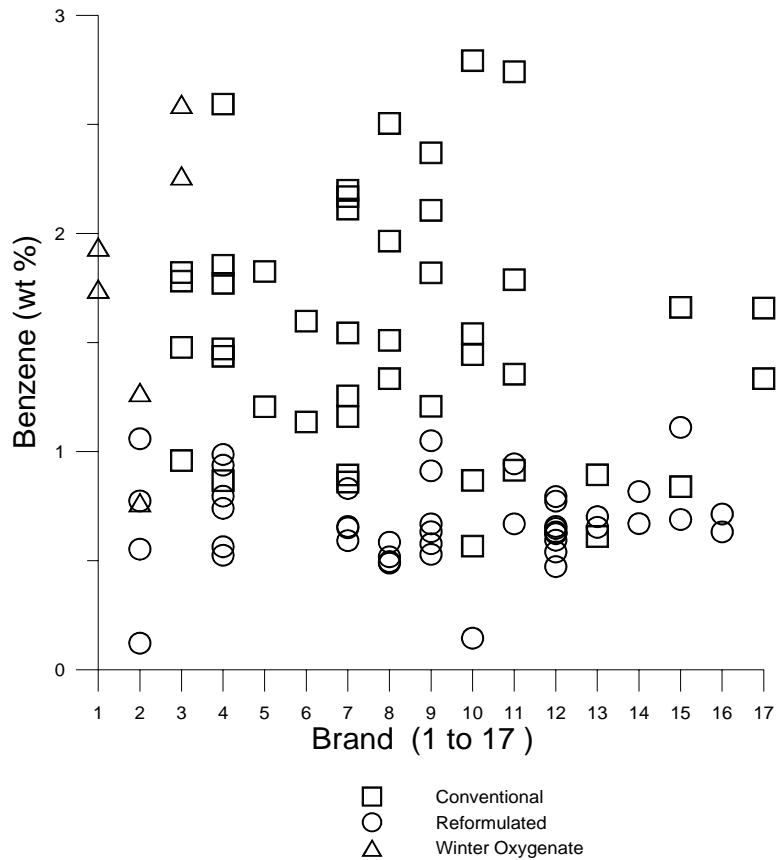


Figure 10 Distribution of benzene by brand.

Similarities and Differences Among Gasolines

Cluster Analysis

The chemical analysis determined the composition of the gasolines in terms of approximately 300 compounds. Since all of these compounds vary from one gasoline to the next, a cluster analysis using all the available concentration data used to determine which gasolines were similar. Figure 11 shows the results of a cluster analysis that included only the samples from Georgia. By including gasolines from only one geographic location, variation due to elevation and differing regulations are eliminated. Differences may remain due to fuel source or vendor, and octane number among others.

The results presented in the figure are drawn as a dendrogram. On the horizontal axis the sample identification is given. These samples are labeled GA for Georgia, 1,2,3, or 4 for the station number, 87 or 93 for the octane number and 1,2,3, or 4 for the replicate number. Notice especially that the samples are separated into two groups differing by octane number, where the 93 octane premium gasolines appear to the left of the center. The vertical axis represents the degree of similarity between two samples. If two samples were identical then the similarity would be 100% and they would be connected along the horizontal axis. For these data, no two samples were exactly the same, so all of the connections occurred at lower levels of similarity. These lower levels are plotted inversely on the vertical axis, so that the further a connection is away from the horizontal axis, the more dissimilar were the samples. Since the samples are all gasolines, there is ultimately a connection made between all the samples, albeit at a low level of similarity.

The cluster analysis grouped the fuels into two clusters which differ by octane number. This result was generated by the clustering algorithm acting only on the compositional data and *not* the octane data. Two samples (GA-1-93-2 and GA-F-1-93-1) differed only by the fill level of the sample vials and are nearly identical. The analysis grouped these two samples closest together of any of the nine at a similarity level of 97.5%. The other three 93 octane samples were included in the grouping as the similarity level dropped to 77.7%. The 87 octane gasolines grouped with only each other although generally at lower similarity levels (64.7%).

Table 8 gives a definition of the clusters in terms of the most prevalent compounds present at 0.50 wt % or more. In either case a different, but overlapping set, of compounds defines the centroid of the cluster. The centroid is analogous to that of a physical object, but here the number of dimensions (chemicals) is much greater than the three dimensions of space. Although defining the cluster, this analysis does not determine which compounds differentiate the clusters.

One noticeable difference between the regular and premium grades was the MTBE content. In the regular grade fuel MTBE was present at an average mass percent of 0.52%, while in the premium fuel MTBE was present at an average mass percent of 3.33%. This indicates that the MTBE may be purposefully present in these Georgia gasolines as an octane enhancer. A

similar result is found in industry data for their Southeast region⁶, Florida and a few other scattered samples (Dickson, 2004a, 2004b).

The increased abundance of chemicals of at least 0.50% weight in Georgia premium grade gasolines compared to regular grades is shown in Table 9. These data show that the compounds that increase are predominantly iso-paraffins, oxygenates and aromatics. The pure compound octane numbers from ASTM (1958) are given in the table. Since many of these compounds are octane enhancers, their octane numbers are generally high (i.e., above that of the regular grade fuel – 87). Similarly, Table 10 shows the compounds that decreased. These include paraffins, oxygenates, aromatics and olefins, generally with lower octane numbers.

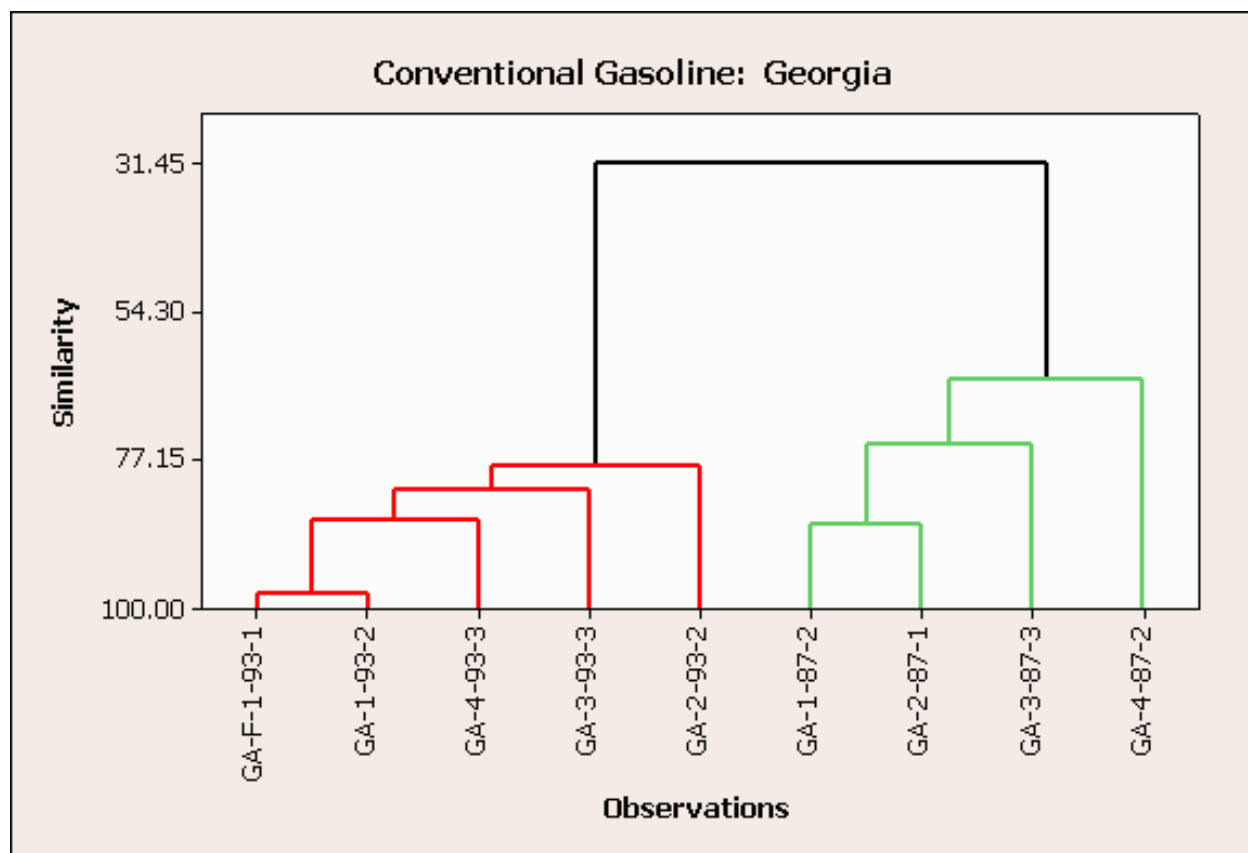


Figure 11 Dendrogram showing the cluster analysis' separation of 93 octane (left) from 87 (right) octane Georgia gasolines.

⁶The region includes North Carolina, South Carolina, Georgia, Tennessee, Alabama, Mississippi, Arkansas, and Louisiana.

Table 8 Cluster centroids in weight % for Georgia 87 octane and 93 octane gasolines defined by all components present at 0.50 wt % or greater.

87 Octane Cluster		93 Octane Cluster	
Toluene	10.3725	Toluene	14.2292
i-Pentane	6.4405	2,2,4-Trimethylpentane	8.4866
m-Xylene	4.3295	i-Pentane	6.108
n-Pentane	4.075	1,2,4-Trimethylbenzene	4.2254
2,2,4-Trimethylpentane	3.477	2,3,4-Trimethylpentane	3.5538
2-Methylpentane	3.3808	Methyl-t-butylether	3.3262
Unidentified	3.2733	m-Xylene	3.246
1,2,4-Trimethylbenzene	2.814	1-Methyl-3-ethylbenzene	2.608
Ethyl-Tert-Butyl-Ethe and Methylcyclopropane	2.534	Unidentified	2.5376
2-Methylhexane	2.493	n-Pentane	2.2486
o-Xylene	2.3658	2-Methylhexane	2.2142
n-Hexane	2.3655	2-Methylpentane	2.1668
Ethylbenzene	2.236	o-Xylene	1.8414
3-Methylpentane	2.2058	Ethylbenzene	1.53
1-Methyl-3-ethylbenzene	2.0488	p-Xylene	1.5
p-Xylene	1.8975	1,3,5-Trimethylbenzene	1.346
3-Methylhexane	1.7058	2,3-Dimethylbutane	1.287
2,3,4-Trimethylpentane	1.6695	1-Methyl-4-ethylbenzene	1.2142
Benzene	1.4538	3-Methylhexane	1.1952
n-Heptane	1.3763	3-Methylpentane	1.1754
n-Butane	1.2463	2,4-Dimethylhexane	1.0838
2-Methylbutene-2	1.1935	2,2,5-Trimethylhexane	1.0666
2,3-Dimethylbutane	1.0755	2,4-Dimethylpentane	1.0514
Methylcyclohexane	0.9983	2,5-Dimethylhexane	1.032
1,3,5-Trimethylbenzene	0.9268	2,2-Dimethylbutane	1.001
1-Methyl-4-ethylbenzene	0.9265	n-Butane	0.9968
t-Pentene-2	0.8038	2,3-Dimethylhexane	0.9952
1-Methyl-2-ethylbenzene	0.7798	1,2,3-Trimethylbenzene	0.9942
3-Methylheptane	0.7228	1-Methyl-2-ethylbenzene	0.9716
1,2,3-Trimethylbenzene	0.7148	n-Hexane	0.903
2,4-Dimethylpentane	0.7038	Ethyl-Tert-Butyl-Ether and Methylcyclopropane	0.8538
n-Propylbenzene	0.701	n-Heptane	0.8476
1,2-Dimethyl-4-ethylbenzene	0.664	Benzene	0.822
n-Octane	0.6523	2-Methylbutene-2	0.8168
2-Methylheptane	0.6495	n-Propylbenzene	0.8078
Cyclohexane	0.6065	1,2-Dimethyl-4-ethylbenzene	0.678
2-Methylbutene-1	0.5715	1-Methyl-3-n-propylbenzene	0.6678
1-Methyl-3-n-propylbenzene	0.5595	t-Pentene-2	0.6048
2,4-Dimethylhexane	0.5475	1,3-Dimethyl-5-ethylbenzene	0.5464
2,3-Dihydroindene	0.5383	TAME	0.5064
2,3-Dimethylhexane	0.5208		
Methyl-t-butylether	0.5188		

Table 9 Compounds that were present in *greater* abundance in Georgia premium than regular grades by at least 0.50 wt %. Octane Number data were taken from ASTM (1958).

Chemical	Pure Chemical Octane Number (R + M)/2	Increase in wt % Regular to Premium
2,2,4-Trimethylpentane	100	5.0096
Toluene	118	3.8567
Methyl-t-butylether	147	2.8074
2,3,4-Trimethylpentane	99.5	1.8843
1,2,4-Trimethylbenzene	136	1.4114
2,2,5-Trimethylhexane	89.5	0.5923
2,2-Dimethylbutane	93	0.5725
1-Methyl-3-ethylbenzene	150	0.5592
2,4-Dimethylhexane	69	0.5363
2,5-Dimethylhexane	--	0.5175

Table 10 Compounds that were present in *lesser* abundance in Georgia premium gasolines than in regular grades by at least 0.50 wt %. Octane Number data were taken from ASTM (1958).

Chemical	Pure Chemical Octane Number (R + M)/2	Decrease in wt% Regular to Premium
n-Pentane	64	-1.8264
Ethyl-Tert-Butyl-Ether and Methylcyclopropane	--	-1.6802
n-Hexane	20.5	-1.4625
2-Methylpentane	80	-1.214
m-Xylene	134.5	-1.0835
3-Methylpentane	83	-1.0304
Unidentified	--	-0.7357
Ethylbenzene	115.5	-0.706
Methylcyclohexane	94	-0.6909
Benzene	94	-0.6318
n-Heptane	0	-0.5287
o-Xylene	111	-0.5244

The similarity of three conventional, low elevation, gasolines were tested by performing a cluster analysis of Georgia, Ohio and Oklahoma gasolines (Figure 12). These samples are designated by state code, station number (1,2,3,4), octane number (87 or 93) and replicate sample number (1,2,3,4). The 87 octanes are clustered with each other on the left most side of Figure 12. The regular grade samples all cluster at 80.05% similarity, while the two main subclusters (samples GA-1-87-2, GA-2-87-3 and GA-3-87-3 form the leftmost cluster) both join at a similarity level of around 85%. The premium grade samples group into three clusters, the

first contains the Georgia samples only. The second and third, while containing one location each, also show lower levels of similarity, 71% and 82%, respectively. This indicates that there is more variability among the premiums than among the regulars, because the premiums fail to cluster together. The poor clustering of the premium samples is partly due to the appreciable MTBE contained in the Georgia premium samples which is absent in those from Ohio and Oklahoma. The MTBE content is the most obvious difference between the fuels but others exist. The Ohio and Oklahoma samples themselves display variation because of their low level of similarity where they cluster. The complete separation by grade in Figure 12, however, shows that the 87 octane and 93 octane gasolines are more similar to each other than fuel of the opposite grade from the same location.

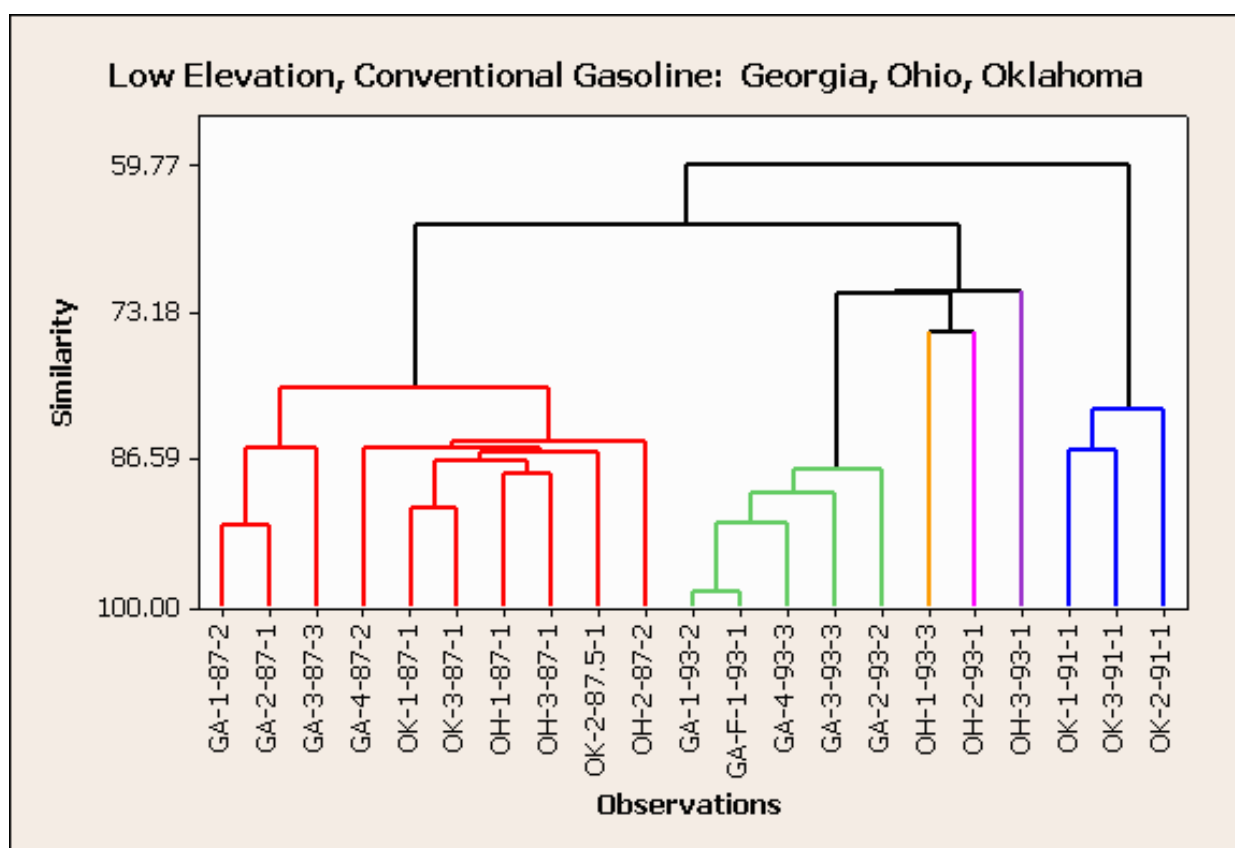


Figure 12 Dendrogram showing relationship between regular and premium gasolines from conventional gasoline areas in Georgia, Ohio and Oklahoma.

The degree of similarity of the premium grades in the entire data set versus that of the regular grades was assessed through a cluster analysis of each location's samples only (similar to the analysis of Georgia gasolines presented in Figure 11). In the majority of locations the regular gasolines were more similar to each other than were the premiums. Table 11 shows the degree of similarity where all samples of a given grade join together. These values are the minimum degree of similarity where all samples of a given grade have joined into a cluster. In

three locales (California, Colorado, and Texas High Plains), there was not good separation between the regular and premium grades. In all other cases the two grades formed distinct clusters.⁷ On average the regular grade gasolines tended to cluster at higher degrees of similarity than did the premiums, meaning that the regular grade fuels were more similar to each other than were the premiums. This result suggests that refiners may be using a variety of approaches to boosting the octane numbers of the premium fuels.

Table 11 Degree of similarity where all fuels of a given grade join in cluster analyses of individual locales (states).

Locale	Degree of Similarity at Complete Joining		Premium Samples Joined to Regular Clusters
	Regular (%)	Premium (%)	
California	55.72	43.87	2 of 3
Colorado	47.98	43.59	2 of 3
Delaware	56.94	82.42	--
Georgia	64.70	77.76	--
Illinois	79.35	61.91	--
Montana	87.44	62.36	1 of 6
New York	60.51	67.72	--
Ohio	66.96	48.37	--
Oklahoma	84.69	80.58	--
Pennsylvania	84.10	21.94	--
Texas High Plains	32.60	32.38	2 of 3
Virginia	38.77	50.04	--
Washington	38.77	37.59	--
<i>Average</i>	<i>64.44</i>	<i>54.66</i>	
<i>Standard Deviation</i>	<i>17.86</i>	<i>19.18</i>	

⁷The regular grade and premium grade samples from one of the Montana gas stations joined the cluster of the opposite grade. It is possible that these samples were mislabeled when sampled or that wrong fuel had been delivered to the tanks at the station.

High elevation samples from Montana and the Texas High Plains are shown in Figure 13. The two states' samples do not join in common clusters. Further, four of the six Texas samples cluster together despite differences in grade, while one premium sample (TX-1-90-2) remains separate from all other samples until the lowest degree of similarity is attained.

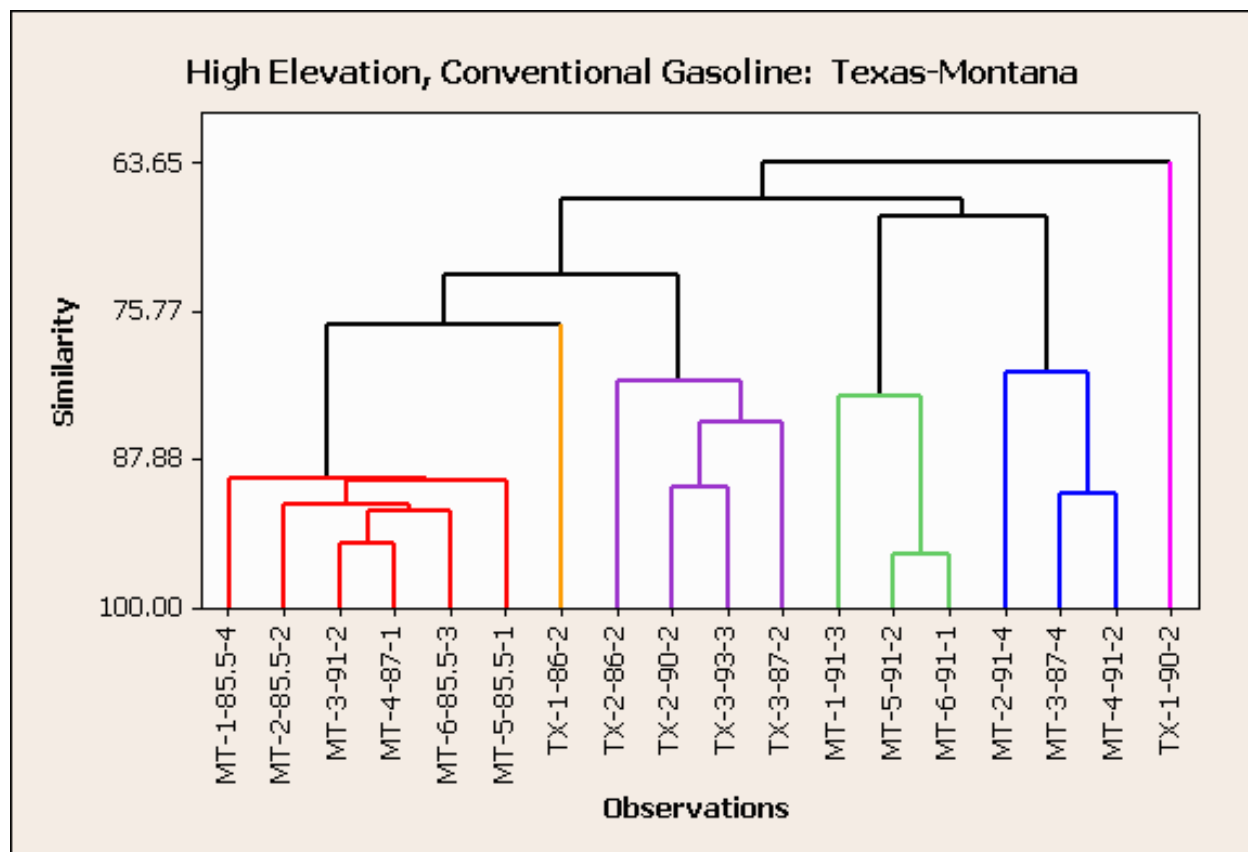


Figure 13 Dendrogram showing relationship between high elevation areas in Lubbock, Texas and Helena, Billings and Great Falls, Montana.

A comparison of reformulated and conventional gasolines is given in Figure 14 by comparing the Georgia gasolines with RFG gasolines from Northern Virginia. Clustering of the Georgia gasolines is the same as shown previously in Figure 11. Four distinct clusters were found in the data. In addition to the two Georgia clusters, the Virginia gasolines formed two grade-dependent clusters. Thus the two different clusters of Virginia gasoline are more similar to each other than they are to the Georgia gasolines. The 93 octane Virginia cluster has a sample (VA-4-93-4) that is unlike the others and only joins the cluster at a low level of similarity (59.0%).

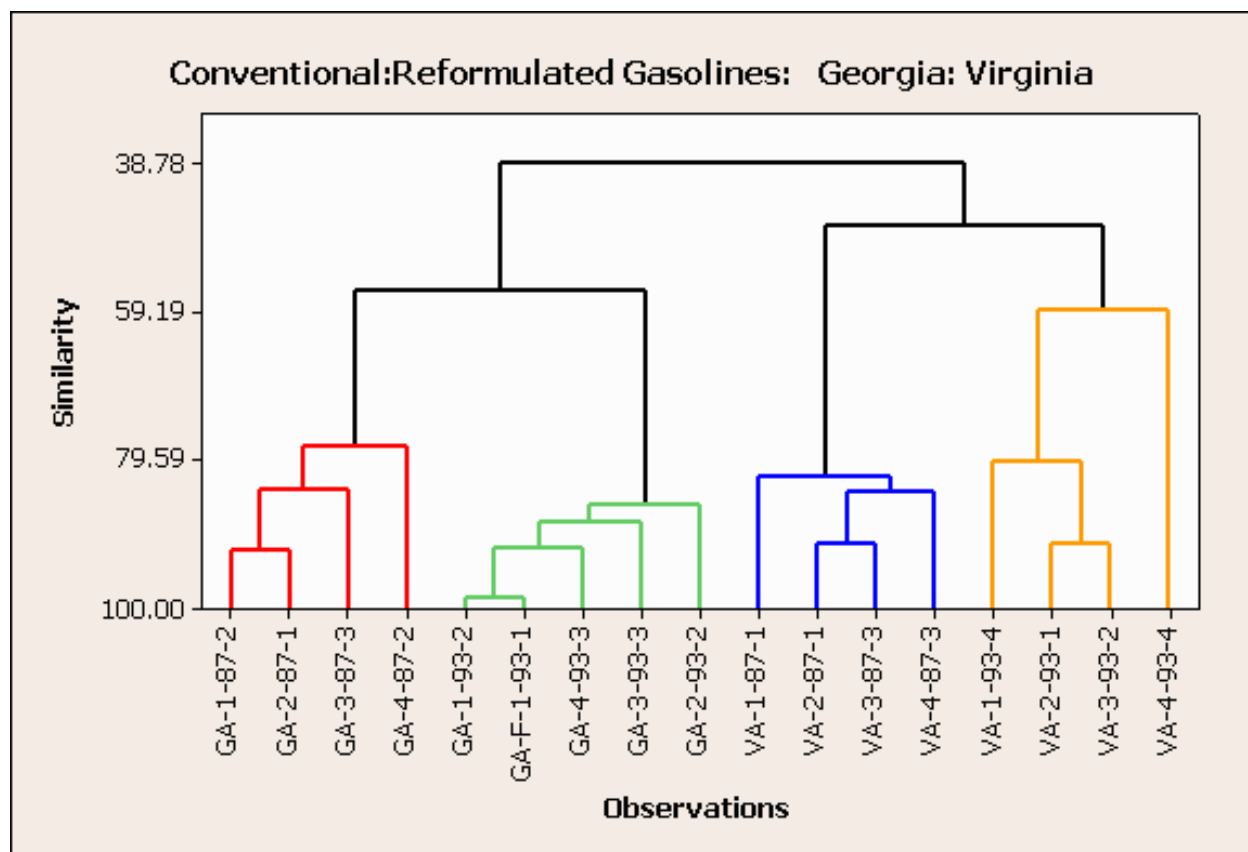


Figure 14 Dendrogram showing relationships between conventional gasolines from Georgia and RFG gasolines from Virginia.

Figure 15 compares gasolines from low elevation, RFG, non-MTBE ban areas in Arlington, Virginia, Dover, Delaware, and Philadelphia, Pennsylvania. These gasolines separate along grade lines first⁸. The regular gasolines form two distinct clusters: one composed of the Virginia gasolines and the other composed of the Dover, Delaware and Philadelphia gasolines. Notably Dover and Philadelphia are close geographically and may consequentially receive similar gasolines. The premium gasolines show similar behavior with a lessened tendency for clustering. The samples VA-4-93-4 and PA-3-93-4 join together but were not sold under the same brand name.

⁸The Delaware 89 octane sample (DE-1-89-2) may represent fuel mislabeled by the retailer.

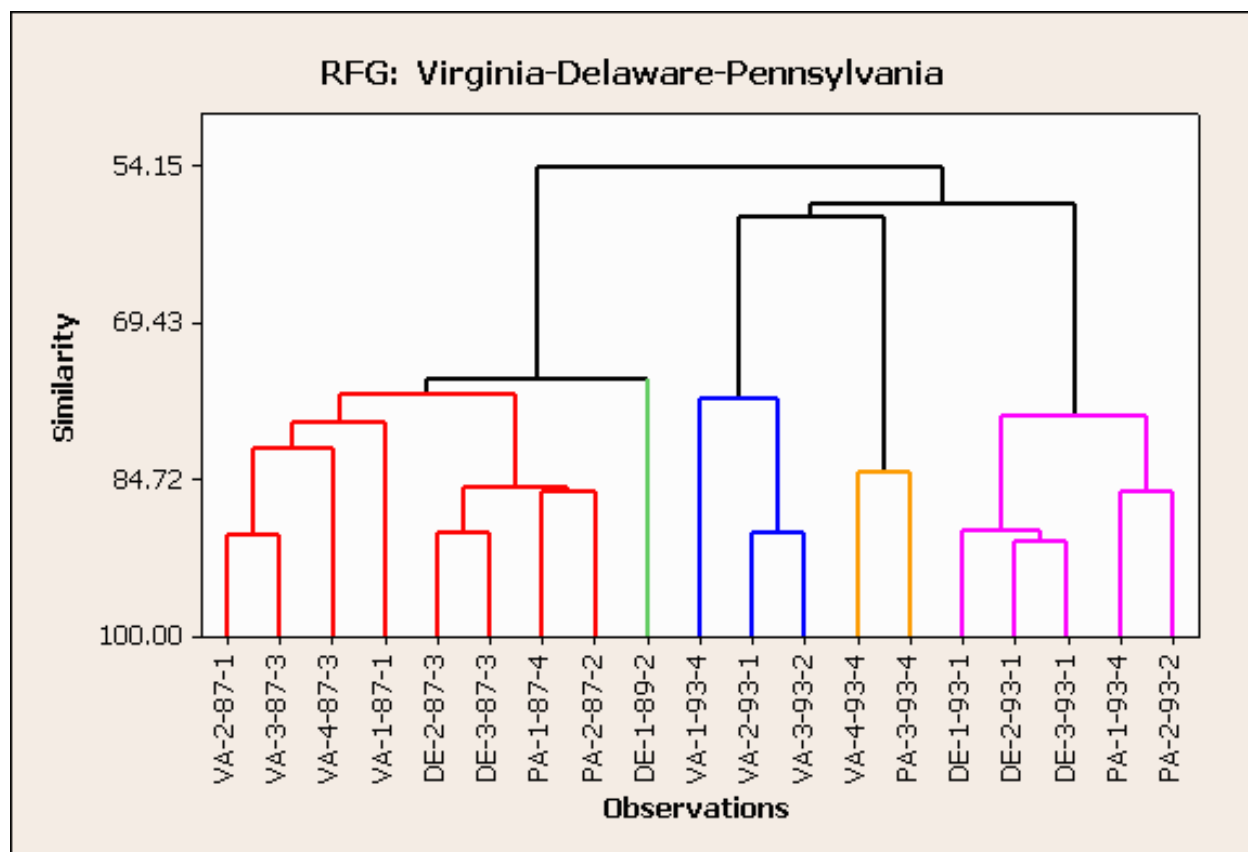


Figure 15 Analysis of three RFG areas without MTBE bans.

Three Federal RFG states in the study have banned MTBE: New York, Illinois and California. To compare gasolines from MTBE-ban states with others, a cluster analysis was performed on the data from Virginia, Pennsylvania and New York. Figure 16 shows that the New York data is separate from the Pennsylvania and Virginia data. Beginning from the left of Figure 16, the New York 87 octane gasolines cluster, followed by the New York 93 octane gasolines. These two clusters join together at a similarity of 71.0%, before they join with the clusters from Virginia and Pennsylvania. The data from these two states tend to cluster with themselves before joining the other state. Even so they join each other at higher levels of similarity, before joining with the corresponding gasoline grades from New York.

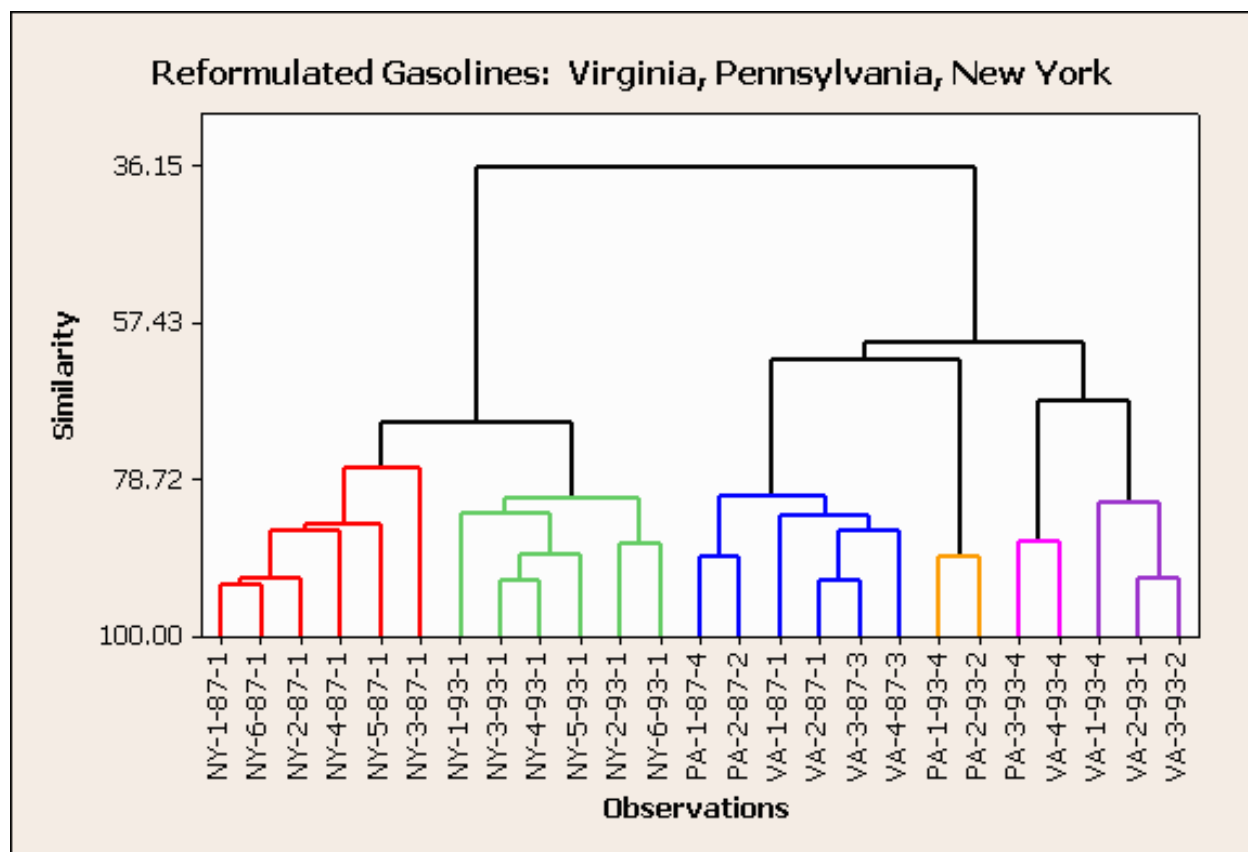


Figure 16 Dendrogram of RFG gasolines from an MTBE-ban state (New York) and two states without bans.

The MTBE-ban states and locales with RFG requirements (state or Federal) are compared in Figure 17. The highest similarity between samples occurred for two California gasolines (CA-3-87-2 and CA-4-87-2). Other California gasolines tended to cluster, but two high octane gasolines were included with the 87 octane fuels. A high degree of similarity occurred with the New York gasolines for both 87 octane and 93 octane gasolines. Some of the Illinois and Colorado samples were grouped with the 87 octane New York gasolines. The Illinois 93 octane samples were similar at a fairly low level (69.5%). Two of the Colorado high octane samples (CO-2-91-1 and CO-3-91-1) grouped with the composited 87 and 93 octane clusters formed largely by the New York samples. Since the Colorado samples were taken from high altitude, and Colorado has specific state RFG requirements, the samples' lack of similarity with the other locales should not be unexpected.

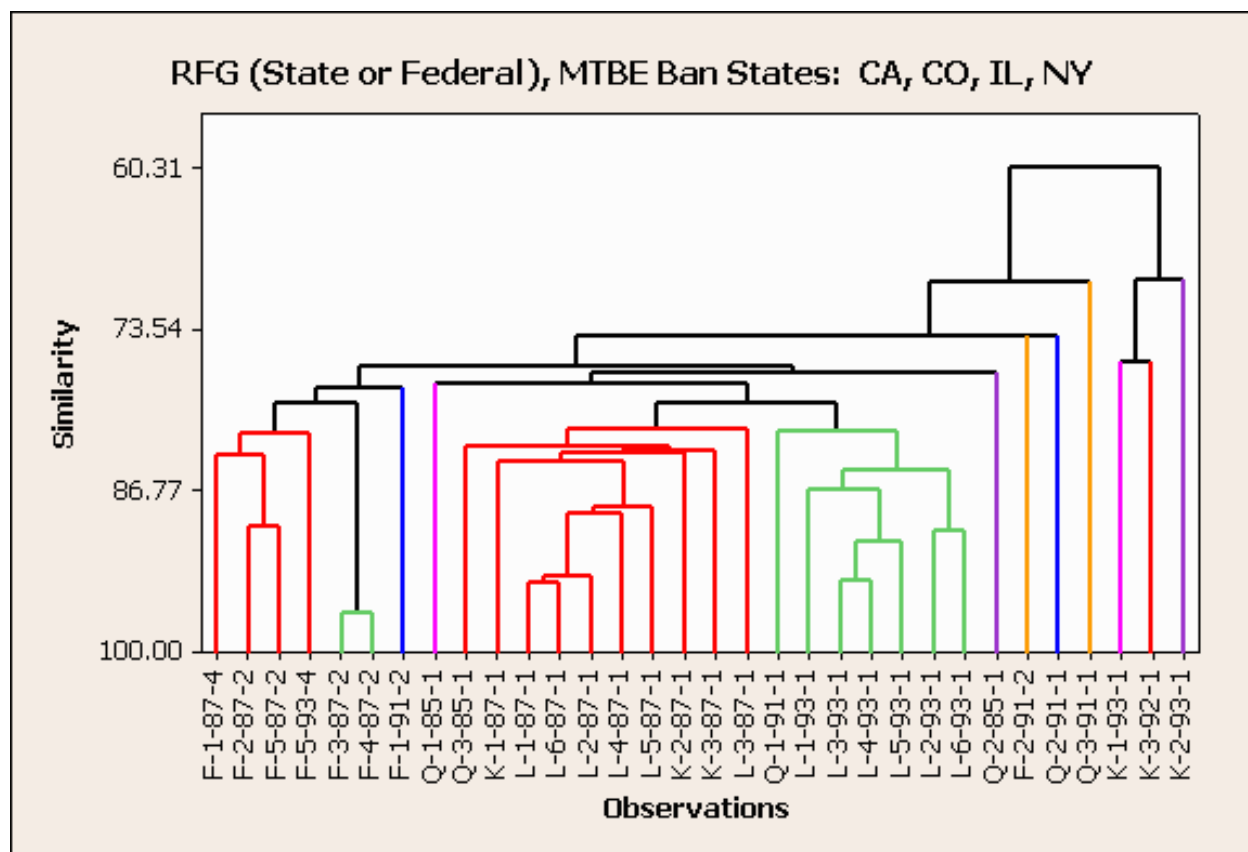


Figure 17 Dendrogram of MTBE-ban states with State or Federal RFG requirements.

Figure 18 shows a cluster analysis of New York and Illinois samples. These both come from states with MTBE bans, low elevation, and RFG requirements. The New York samples form a cluster, that at lowered levels of similarity, are joined by the Illinois samples. The premium gasolines from New York cluster with these regular gasolines at *higher* levels of similarity than they do with the Illinois premiums. Again this result suggests a low level of similarity among the premium grade samples.

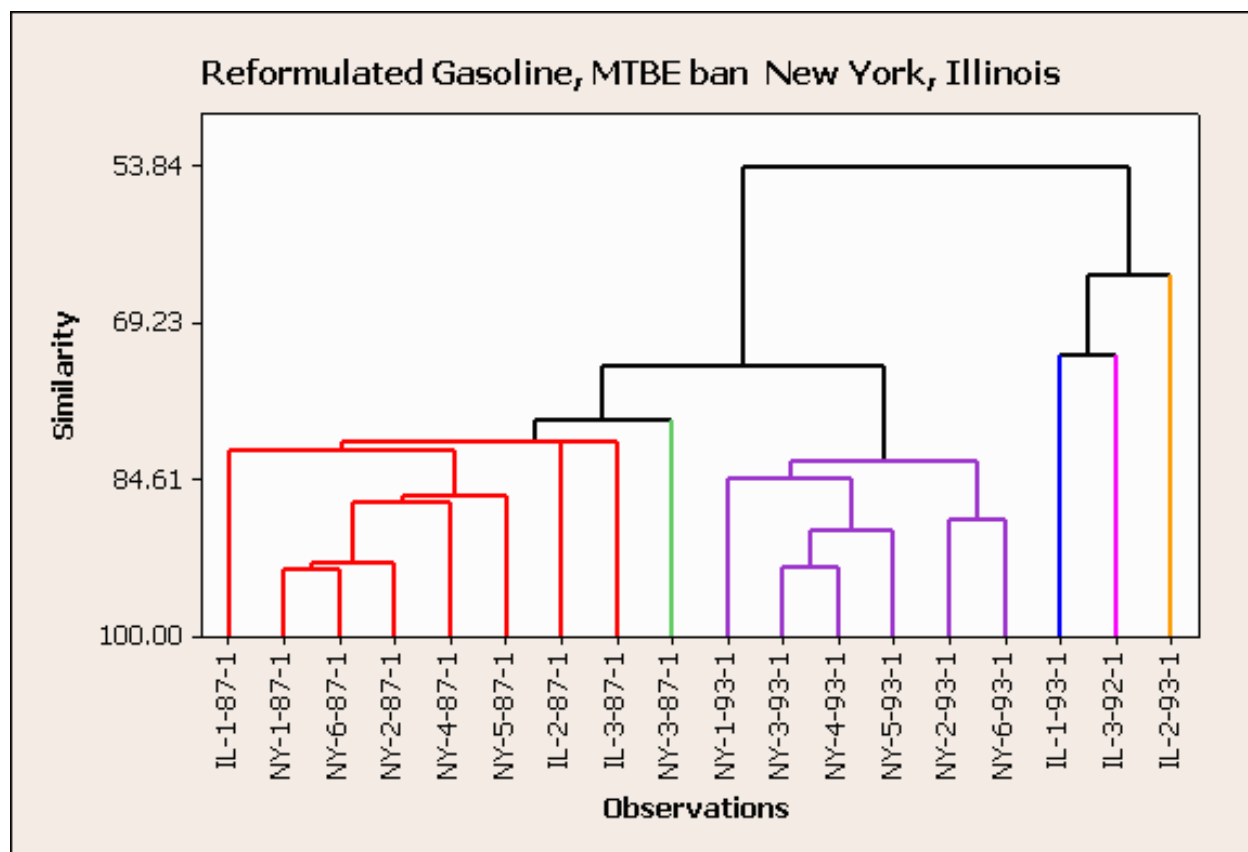


Figure 18 Dendrogram showing relationship between two MTBE ban, RFG gasoline areas.

Principal Components Analysis

The basic goal of principal component analysis (PCA) is to reduce the dimensionality of a multivariate data set consisting of a large number of variables, while capturing as much of the variation as possible in the data set. PCA is a mathematical technique that transforms a number of potentially correlated variables into a small number of uncorrelated variables called principal components (PC). The first PC accounts for as much of the variability in the data as possible, and each succeeding PC accounts for as much of the remaining variability as possible (Kendall 1975, Jolliffe 2002). Further background information on PCA is given in an Appendix (page 79). The PCA performed on the gasoline data served two purposes. First it complimented the results from the cluster analysis by indicating which groups of samples were similar to each other. Second, it gave information on the components that had the biggest influence on the groupings.

Results of the PCA on the individual gasoline components are presented in Table 12. Only 8 PC were required to account for 95 percent of the variation in the gasoline samples. The total variance estimate of the gasoline components was found to be 87.53. The first two PC accounted for almost 60 percent of the variation of the gasoline components. A plot of the cumulative proportion of variation explained by PC number is given in Figure 19. The figure

shows that between 99 and 100 percent of the variation in the gasoline components can be explained by 16 and 33 PC, respectively. A scree diagram is a plot of the variance of each PC against the PC number, p (the rank importance of the PC, see page 79 for more information). Figure 20 contains the scree plot of the variance estimates by PC number which shows that the variance estimates decrease rapidly until approximately the tenth PC. Beyond this point there was little change in the variance estimates.

Table 12 Statistics of the first eight PC from the PCA of the individual gasoline components including variance, proportion of the total variance estimate, and the cumulative proportion of variation explained.

PC	Variance	Proportion	Cumulative
1	29.493	0.337	0.337
2	20.652	0.236	0.573
3	16.408	0.188	0.760
4	8.574	0.098	0.858
5	3.189	0.036	0.895
6	2.522	0.029	0.924
7	1.376	0.016	0.939
8	1.169	0.013	0.953

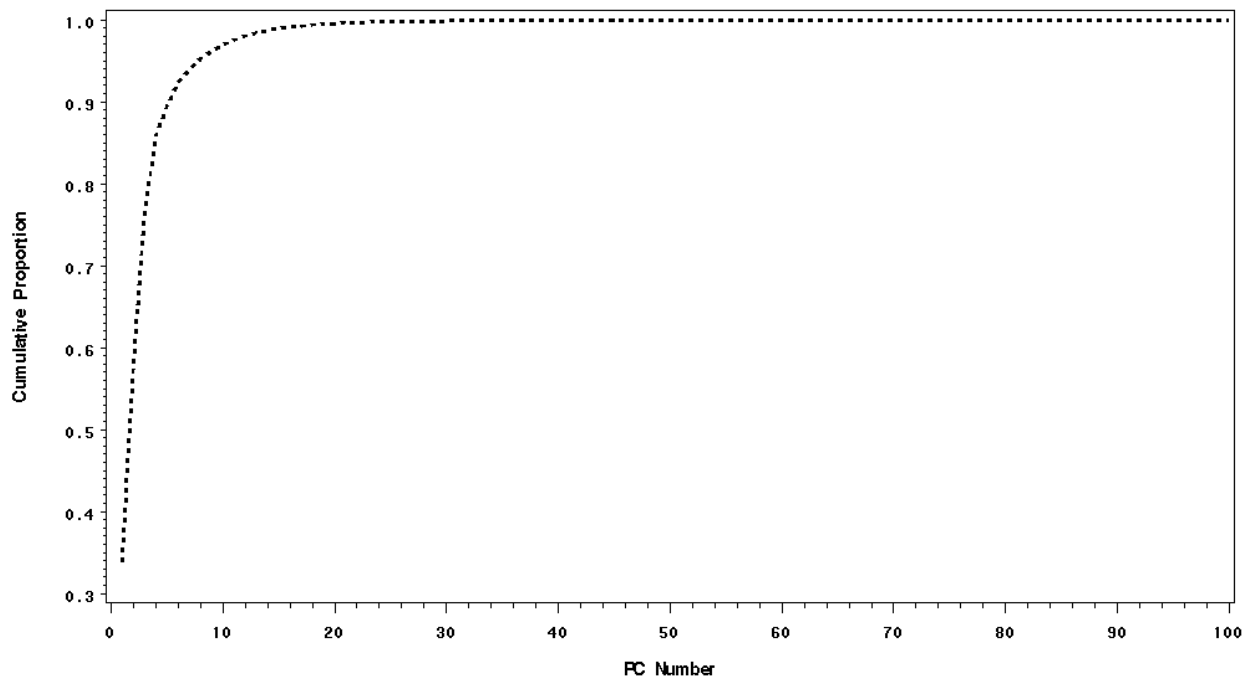


Figure 19 Plot of the cumulative proportion of variation explained by PC number.

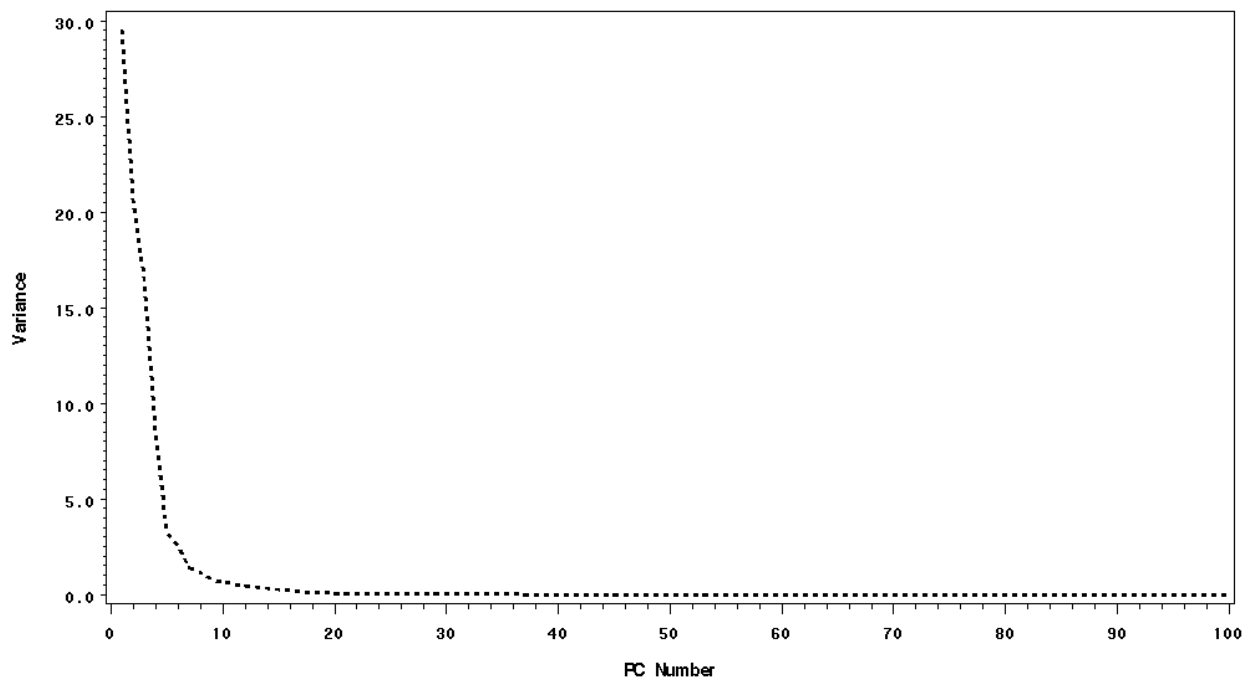


Figure 20 Scree graph of the variance estimate by PC number for the individual gasoline component data.

The largest loadings from the eight PC indicate which components contribute the most to the variation (Table 13). Table 13 shows that the gasoline component corresponding to the largest absolute loading in the first PC was 2,2,4-Trimethylpentane. This component also had the largest standard deviation (rank =1) among all of the gasoline components analyzed. From Table 13 we can also see that those gasoline components with the largest coefficients consistently had the largest standard deviations. Thus, the eight gasoline components presented in Table 13, represent those components which contributed the most to the variation among the gasoline samples. As discussed below, the first two principal components were dominated by 2,2,4-Trimethylpentane (isooctane) and MTBE, representing the influence of octane number and oxygenated additives, respectively.

Table 13 Gasoline components with the largest loadings in the first eight principal components including class, mean, standard deviation, and the rank of the standard deviation (from largest (1) to smallest(8)).

PC	Gasoline Component	Class	Mean	Std. Dev.	Rank Std. Dev.
1	2,2,4-Trimethylpentane	Iso-Paraffins	5.62	4.60	1
2	Methyl-t-butylether	Oxygenates	2.31	4.07	2
3	Toluene	Mono-Aromatics	8.15	3.35	4
4	Ethanol	Oxygenates	2.28	3.36	3
5	2-Methylhexane	Iso-Paraffins	3.56	2.34	5
6	Unidentified1		3.93	1.60	7
7	n-Butane	Paraffin	2.50	1.11	10
8	n-Pentane	Paraffin	2.74	1.40	8

A plot of the first two principal components labeled by state is given in Figure 21. From the figure it can be seen that marked differences exist among the states, with the data falling into roughly five clusters. Samples from Virginia, Pennsylvania/New Jersey, and Delaware were concentrated at the bottom of the graph. These fuels are all low elevation RFGs relying on MTBE for their oxygen content. Georgia premiums lay in the middle of the graph. As previously noted, these conventional gasolines contained MTBE presumably as an octane enhancer. At a level of the 2nd principal component of about 0 to 2, a large group of samples form a thin band. These included the majority of the conventional gasolines. At a level of 3 to 5, a thin band was distinguished by the four locations with RFG requirements and MTBE bans. These gasolines shared the common characteristic of low MTBE concentration and ethanol use to meet their oxygen requirement. The last separate cluster consisted of gasolines from Colorado, that were under an oxygenate requirement. The 2nd principal component can be identified with the two dominant components that supply oxygen in the fuels: MTBE which is at a maximum at the bottom and ethanol which is at its maximum at the top of Figure 21.

Marked differences appeared between the octane levels (Figure 22). With a few exceptions, the octane numbers fell into two classes, 87.5 or lower and 90 or above. The regular gasolines (87.5 or below) were grouped to the right of the value of about 3 on the horizontal axis, while the premium gasolines (90 or above) levels fell to the left of this value. The first principal component was most highly influenced along its negative axis (toward high octane fuels) by 2,2,4-Trimethylpentane (isooctane) and other iso-paraffins. In the positive direction (toward low octane fuels) it was influenced most strongly by paraffins and a variety of compounds of the other classes. Isooctane had the strongest influence of any chemical on this principal component. The mean values 2.78 and 8.46 % were for the 87.5 or lower and 90 or above octane groups, respectively. The dominance of isooctane explains the ability of this principal component to separate the octane classes. Isooctane is neither the only chemical that influences the octane number, nor the principal component. Thus there was not a direct relationship between the principal component and the octane numbers.

A plot of the first two principal components labeled by gasoline brand is given in Figure 23. No obvious patterns can be detected among the 17 brands under which the gasolines were sold.

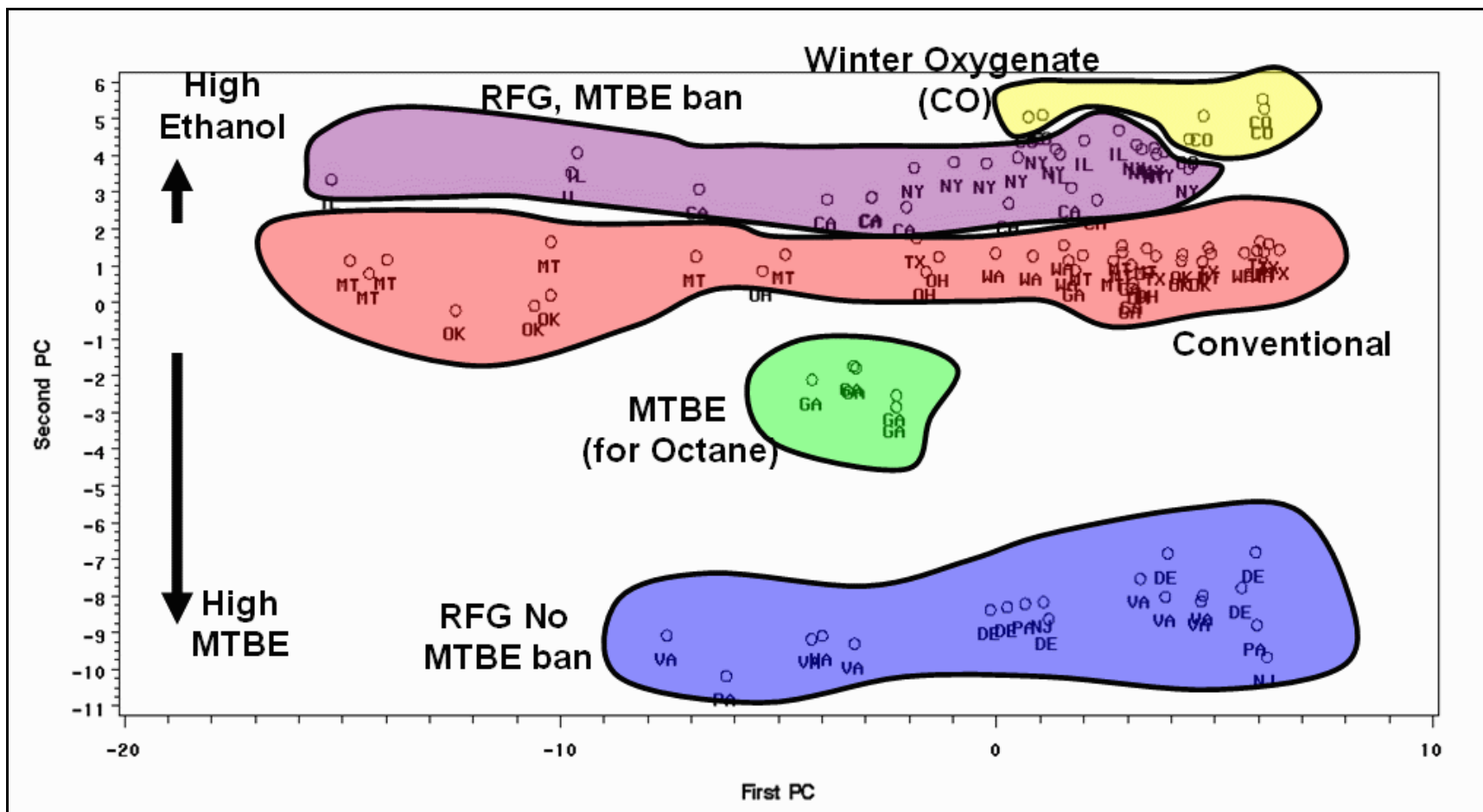


Figure 21 Principal Components plot of the 96 gasoline samples with respect to the first two principal components labeled by state.

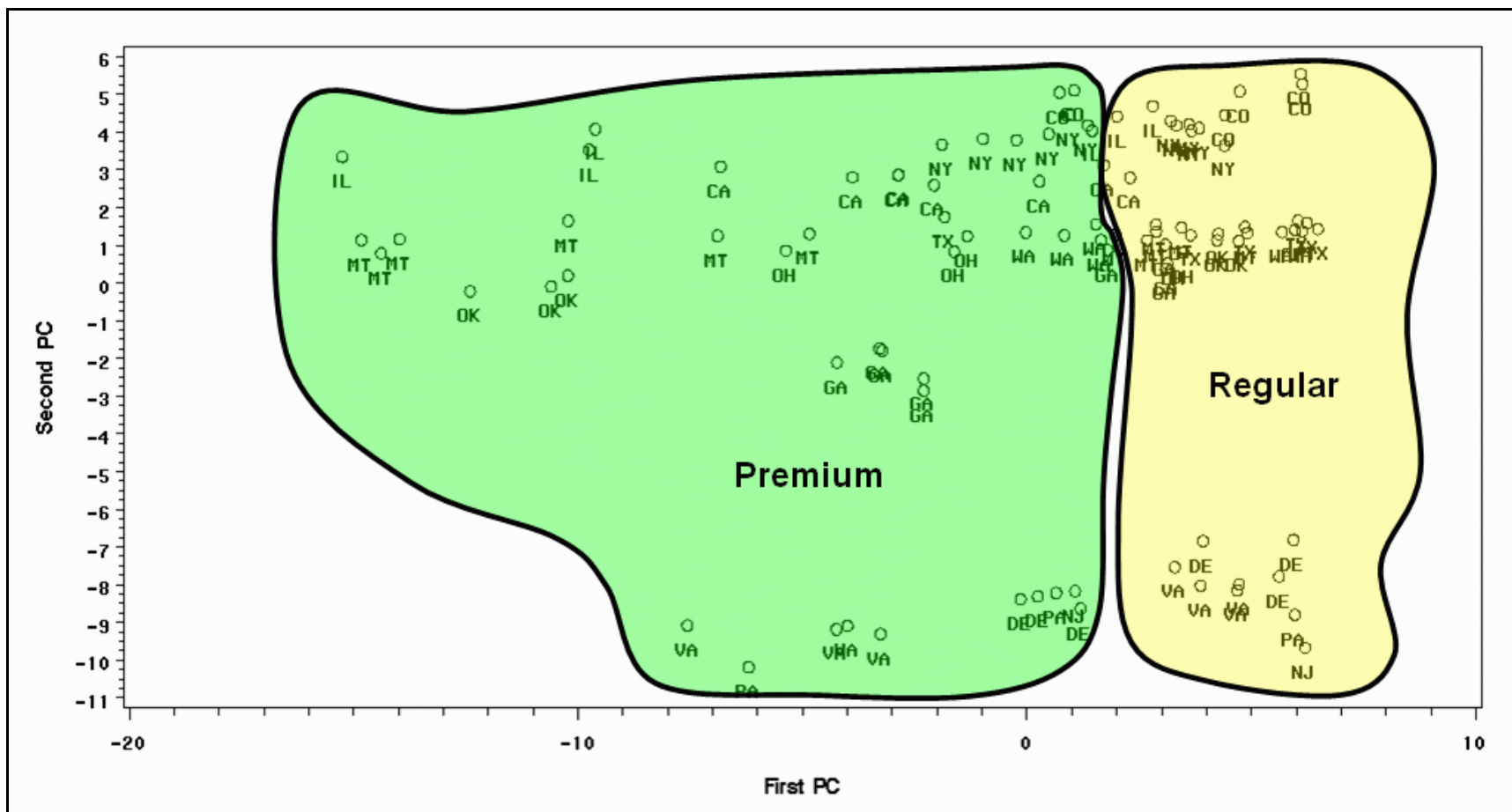


Figure 22 Principal Components plot of the 96 gasoline samples with respect to the first two principal components labeled by octane number.

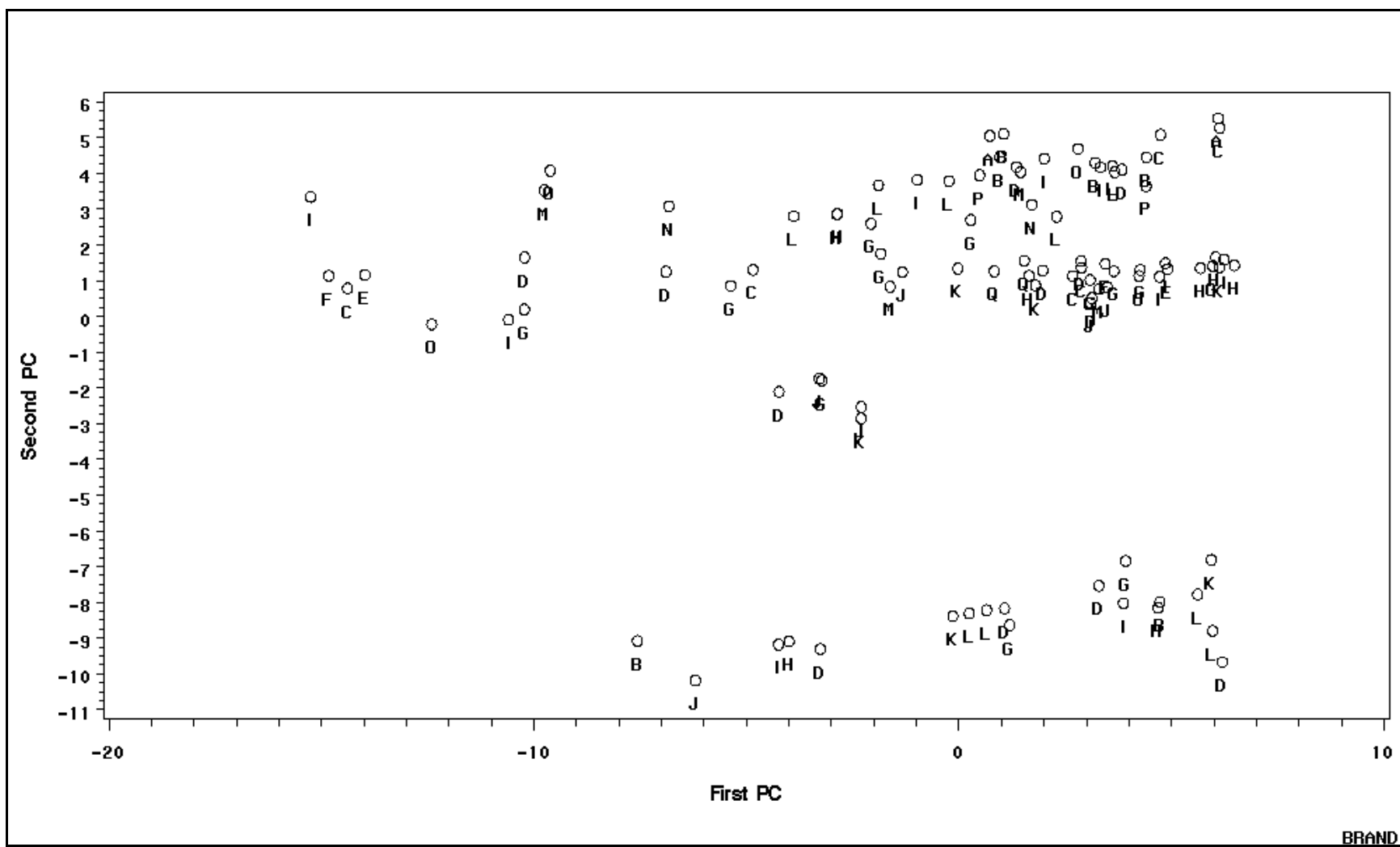


Figure 23 Plot of the 96 gasoline samples with respect to the first two principal components by gasoline brand.

Discriminant Analysis

A step-wise discriminant analysis was used to determine which chemical components were most powerful in distinguishing among the samples (see Rencher, 1995). The analysis tests the statistical level of significance of each variable and determines if it contributes to discriminating among the samples. The results of this analysis are shown in Table 14. The significance level was high for RFG/Conventional gasoline, MTBE ban, grade, elevation and altitude. It was low for all factors taken together and brand. Thus gasolines having the same values for the characteristics and a high significance level are expected to separate well based on composition. Two clear compositional indications emerged: for distinguishing an MTBE ban gasoline, the most important chemical was ethanol. This follows from the substitution of ethanol for MTBE in CA, CO, IL and NY. Ethanol and MTBE were most important for distinguishing RFG from conventional gasolines, clearly because of their usage as oxygenates.

Table 14 Discriminant analysis results.

Group Tested	Significance Level	Number of Components Required	Five Most Significant Components
All	0.59	18	Methyl-t-butylether Ethanol n-Hexane 3t-Ethylmethylcyclopentane n-Pentadecane
RFG/Conventional gasoline	0.99	19	Ethanol Methyl-t-butylether 1c,2t,4-Trimethylcyclopentane C9-Olefin4 o-Xylene
MTBE Ban (y/n)	0.96	20	Ethanol n-Pentadecane 3t-Ethylmethylcyclopentane C10-Isoparaffin2 3,4-Dimethylpentene-1
Brand (17)	0.40	22	3,5,5-Trimethylhexene-1 1-Methyl-3-n-propylbenzene 1t-M-2-(4-MP)cyclopentane n-Pentadecane 2,3,3-Trimethylbutene-1
Grade (<89, >90)	0.95	23	Ethyl-Tert-Butyl-Ether and Methylcyclopropane 4-Methylnonane n-Hexane 2-Methylindan 1t,3-Pentadiene
Altitude (H/L)	0.94	23	C11-Isoparaffin2 i-Propylbenzene 2,2-Dimethylpropane Benzene 2-Methylbutene-1

Representative Gasolines

Figure 24 shows a decision tree that represents four characteristics of the fuels: grade, sale elevation, MTBE ban status, and legal requirements. The latter include the federal reformulated gasoline mandate, state requirements (as the California RFG requirement, or Colorado winter oxygenate requirement) and, alternatively, conventional gasoline. These provide a potential means for dividing or grouping samples. For a given grade of gasoline there are 12 branches which may represent 12 or more types of gasoline because of the number of state requirements. Seven of the branches were represented in this study. In all cases more data would be desirable to increase the number of samples in each category.

Division of the data by the characteristics is supported to some degree by the statistical analyses. These provide a scientific basis for what appear to be logical groupings. The discriminant analysis supported division by grade, elevation, MTBE ban and RFG requirement as these were found to be factors that provided a high degree of reliability of discriminating between samples. The principal components analysis supported division by grade and combined MTBE ban status and RFG requirements. The cluster analysis generally supports division by grade, RFG requirement, and MTBE ban status, though not elevation. Because it was performed on subsets of the data, the cluster analysis showed that the premium gasolines were more variable than the regular gasolines.

Figure 24 indicates where the data are likely to be best grouped together. At the ends of the decision tree, parenthesis around state names indicate that the cluster analysis indicated similarity. All of these are not clear cut, and other judgements could be made to separate or join data together. One argument in favor of grouping data together at the end of each branch of the decision tree is that the variability will be maximized by this procedure. Risk assessments relying on typical data, such as these, should include the maximum range of variability that might occur.

Grade	Elevation	MTBE Ban Status	Legal Requirements
Premium	> 3000 ft	ban	Federal RFG
			State Requirement CO
			Conventional
		no ban	Federal RFG
			State Requirement
			Conventional (TX), (MT)
	< 3000 ft	ban	Federal RFG (NY), (IL)
			State Requirement CA
			Conventional WA
		no ban	Federal RFG (VA), (DE, PA)
			State Requirement
			Conventional (GA), (OH), (OK)
Regular	> 3000 ft	ban	Federal RFG
			State Requirement CO
			Conventional
		no ban	Federal RFG
			State Requirement
			Conventional (TX), (MT)
	< 3000 ft	ban	Federal RFG (NY, IL)
			State Requirement CA
			Conventional WA
		no ban	Federal RFG (VA, DE, PA)
			State Requirement
			Conventional (GA, OH, OK)

Figure 24 Decision tree for gasolines typical of the data collected in this study. The parenthesis indicate where the cluster analysis indicated that samples from the various states are similar enough to group together.

Cluster Analysis

The degree-of-similarity of the representative gasolines was tested by performing a cluster analysis. Figure 25 shows the results from averaging all samples at each branch of the decision tree. For the maximum differentiation between the representative gasolines, ideally the clustering would be at low levels of similarity. The identification of these data are: RFG – N/Y/Y-State, MTBE ban – Y/N, grade R/P, elevation H/L. A few clear indications emerged from this figure. Beginning on the right hand side, the low elevation, RFG, grades join at a low level of similarity (Y-N-R-L and Y-N-P-L). Some RFG gasolines do not cluster well with others: YS-Y-P-H and Y-Y-P-L. The remaining categories showed a fairly low level of similarity for the most part, with the exception of the N-N-R-H and N-Y-R-L samples which join at a level of similarity (85%) that was judged to be favorable for individual samples.

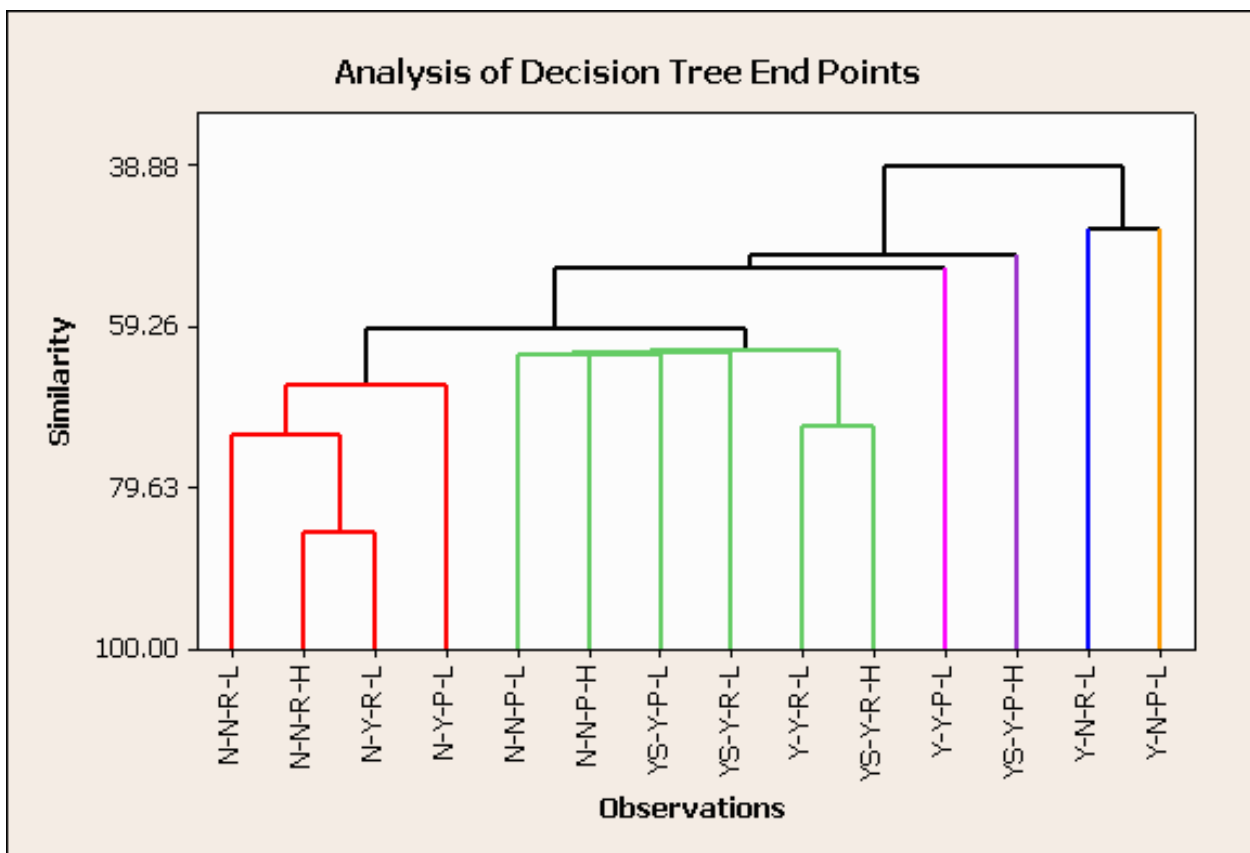


Figure 25 Cluster analysis performed on the representative gasolines.

Predicted Environmental Impacts of the Study Fuels on Water and Air

From an environmental contamination perspective, fuel composition is most relevant for soil contamination where the fuel itself is contained within the sample. Ground water and soil gas contamination are equally important, but the relevance of each compound depends on its abundance in the fuel and its properties.

Raoult's Law partitioning relationships were used to determine the equilibrium concentration of each chemical in air and water at 15 °C (Schwarzenbach et al., 2003).

$$S_{e_i} = x_i S_i \quad (1)$$

where S_{e_i} is the effective solubility of chemical i resulting from the gasoline mixture [mg/L], x_i is the mole fraction of chemical i in the mixture [moles/moles], and S_i is the solubility of chemical i [mg/L]. Similarly the effective vapor pressure due to the presence of the chemical in the mixture are given by Raoult's Law

$$P_{e_i} = x_i VP_i \quad (2)$$

where P_{e_i} is the partial pressure of chemical i resulting from the gasoline mixture [atm], x_i is the mole fraction of chemical i in the mixture [moles/moles], and VP_i is the vapor pressure of chemical i [atm].

Because solubility and vapor pressure data are not available for all hydrocarbons at all temperatures of interest and because there exists considerable variability in literature values, the SPARC Performs Automated Reasoning in Chemistry (SPARC) calculator was used to estimate all properties used in the analysis. SPARC estimates are generally viewed to be within one-third to one-half log unit of true values. Thus the SPARC estimates may differ from commonly used estimates of some parameters. SPARC recognized 292 of the 296 chemicals with CAS numbers in the data set.⁹ The disadvantages of SPARC were outweighed by the advantage of having a consistent set of parameters for all chemicals at all temperatures of interest (0 °C to 25 °C). The mole fractions, x_i , were calculated from the mass fractions, m_i , by

$$x_i = \frac{\frac{m_i}{gmw_i}}{\sum \frac{m_j}{gmw_j}} \quad (3)$$

⁹In addition to the specific chemicals, 5 groups of unknowns and 30 groups of generalized composition were reported.

and using the gram-molecular weights, gmw_i [g/mole]. Equation 3 was used by assuming a sample size of 1g of gasoline. From this mass the number of moles of each identifiable chemical was determined, summed and used to calculate the mole fractions. Chemicals that were unidentified were omitted from the calculation, so the results gave an approximation of the true mole fractions.

Gas phase concentrations, C_a , that would be associated directly with the fuel were calculated from equations 2 and 3, and the ideal gas law:

$$C_a = \frac{P_{e_i}}{R T} gmw_i \quad (4)$$

where R is the gas constant, 8.205×10^{-5} atm - m^3 /mole °K, and T is the temperature in Kelvin.

Predicted Aqueous Concentrations

Figure 26 shows a plot of the effective solubility of each chemical, labeled by group (oxygenate, BTEX, parafins, etc), and plotted against the weight % in the sample. The data were the average conventional gasoline, no MTBE ban, low elevation, regular grade samples. Only solubilities above 0.001 mg/L were plotted as any lower values would be below typical detection limits. The plotting position on this graph indicates the partitioning behavior of the compound. For example, compounds plotting at the lower right (high weight % and low effective solubility) are present in large quantities in the fuel but have little impact on water quality. Conversely, the plotting in the upper right are compounds with high effective solubility and abundance in the fuel. These are the major components of fuel that do potentially impact water quality.

The highest effective solubilities (10 mg/L to 100 mg/L) occurred for oxygenates and BTEX (Figure 26). In the next lower group (1 mg/L to 10 mg/L) were representatives of all the remaining groups. Many hydrocarbons plot lower in effective solubility (less than 1 mg/L) and also occur at low weight percent.

Similar results for the average no-MTBE-ban, low-elevation, regular-grade, reformulated gasoline are shown in Figure 27. Here MTBE had the highest effective solubility, again followed by BTEX and oxygenates in the 10 mg/L to 100 mg/L range. Table 15 shows predicted concentration increases greater than 0.05 mg/L between conventional gasoline and RFG. The compounds with the highest increases are all oxygenates, followed by naphthenes and olefins. Table 16 shows the corresponding decreases in concentration. The compounds with the largest decreases are BTEX, and mono-aromatics, followed by a few paraffins, an iso-paraffin and diolefins.

If an MTBE ban were introduced, gasoline composition might be similar to the data obtained from the Illinois samples. The average effective solubility of chemicals in the average

Illinois sample are plotted in Figure 28. The highest effective solubility was found for ethanol, followed by a similar pattern for the other fuels. For this potential change in fuel, there are a smaller number of chemicals with increased effective solubility (Table 17), including the alcohols and benzene along with two paraffins and an iso-paraffin. Compounds with decreased effective solubility (Table 18) are the ether oxygenates, BTEX other than benzene, and a suite of other compounds including n-olefins, iso-olefins and other compounds.

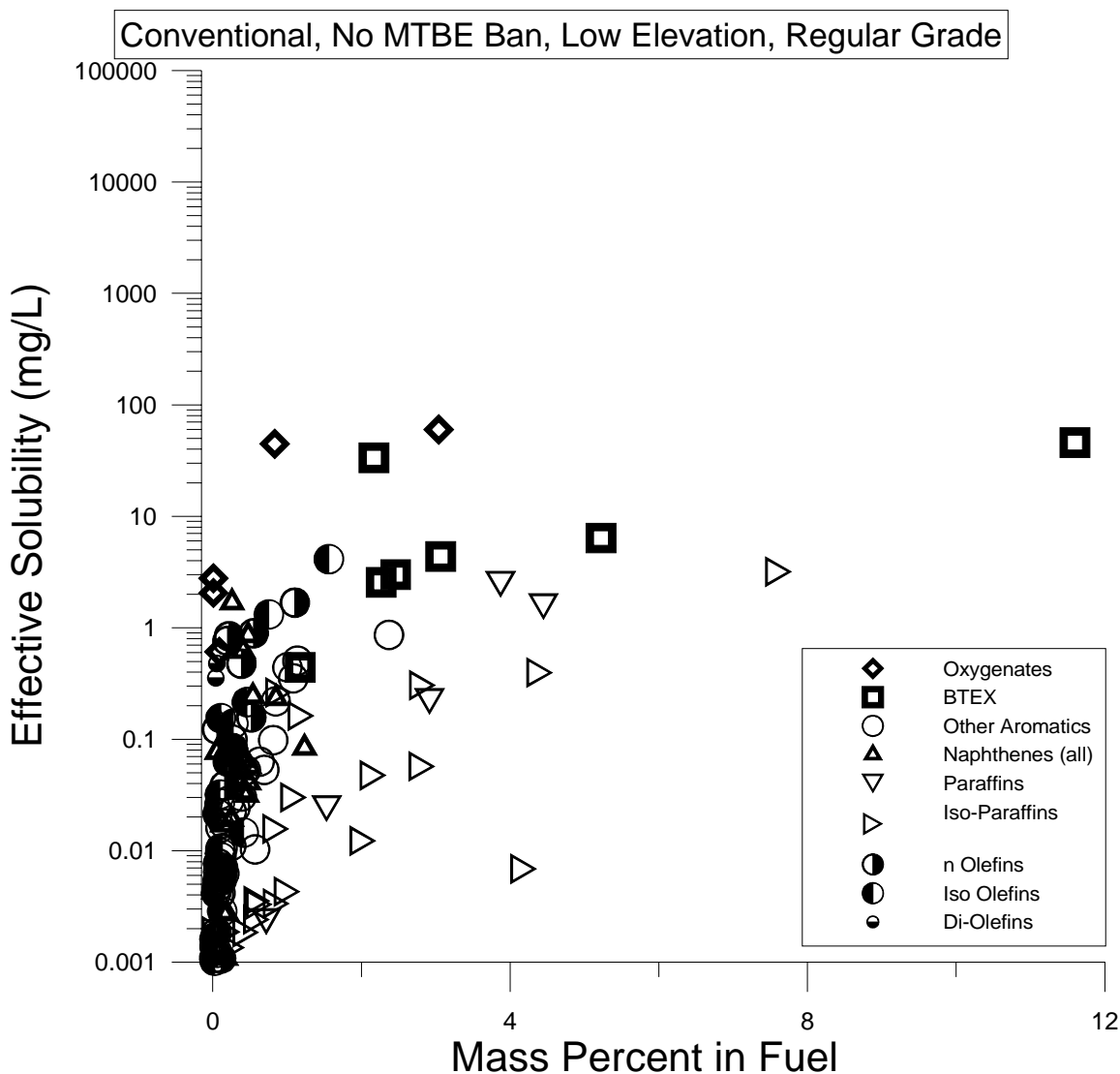


Figure 26 Effective solubility of hydrocarbon groups and oxygenates for the average conventional, low-elevation, no-MTBE-ban, regular gasolines.

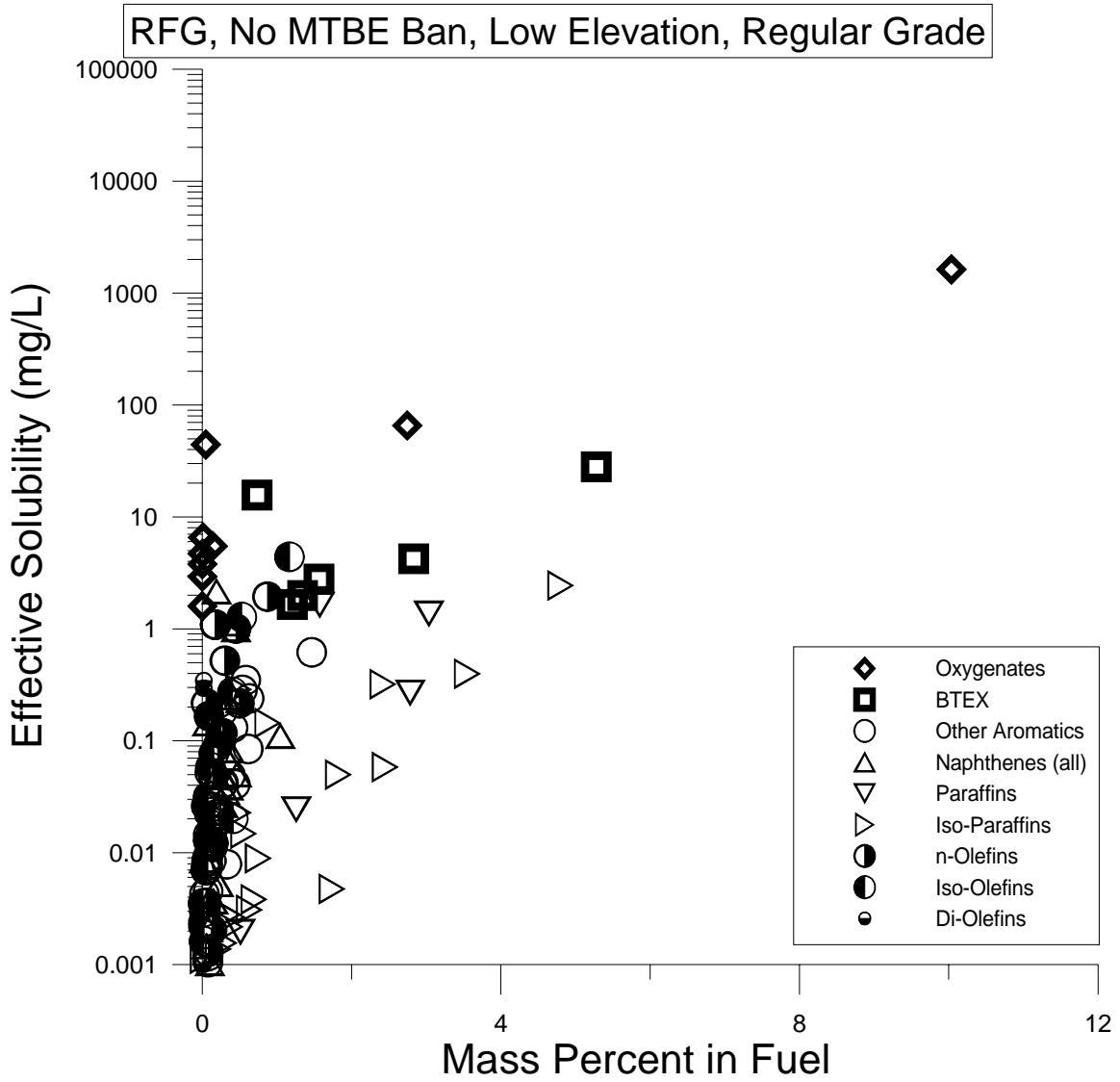


Figure 27 Effective solubility of hydrocarbon groups and oxygenates for the average reformulated, low-elevation, no-MTBE-ban regular gasolines.

Table 15 Predicted *higher* aqueous constituent concentrations (>0.05 mg/L) of RFG (regular grade, no-MTBE-ban, low-elevation) in comparison to conventional gasoline.

Class	Chemical	Change in Concentration (mg/L)
Oxygenates	Methyl-t-butylether	1580
Oxygenates	t-Butanol	42.2
Oxygenates	TAME	4.89
Oxygenates	2-Butanol	4.70
Oxygenates	n-Butanol	3.80
Oxygenates	i-Propanol	3.78
Oxygenates	n-Propanol	2.92
Oxygenates	i-Butanol	1.59
Naphtheno-Olefins	1-Methylcyclopentene	0.50
Naphtheno-Olefins	Cyclopentene	0.48
n-Olefins	t-Butene-2	0.33
Iso-Olefins	2-Methylbutene-2	0.27
n-Olefins	t-Pentene-2	0.26
n-Olefins	c-Butene-2	0.25
Mono-Naphthenes	Cyclopentane	0.15
n-Olefins	c-Pentene-2	0.11
n-Olefins	Butene-1	0.09
Naphtheno-Olefins	Cyclohexene	0.07
n-Olefins	t-Hexene-2	0.06
Iso-Olefins	2-Methylpentene-2	0.06

Table 16 Predicted *lower* aqueous constituent concentrations (>0.05 mg/L) of RFG (regular grade, no MTBE ban, low elevation) in comparison to conventional gasoline.

Class	Chemical	Change in Concentration (mg/L)
Mono-Aromatics	Toluene	-17.80
Mono-Aromatics	Benzene	-17.80
Mono-Aromatics	m-Xylene	-2.16
Mono-Aromatics	o-Xylene	-1.58
Mono-Aromatics	Ethylbenzene	-1.00
Mono-Aromatics	p-Xylene	-0.90
Paraffin	n-Butane	-0.83
Iso-Paraffin	i-Pentane	-0.74
Mono-Aromatics	1,2,4-Trimethylbenzene	-0.52
Mono-Aromatics	1-Methyl-3-ethylbenzene	-0.25
Paraffin	n-Pentane	-0.19
Mono-Aromatics	1,2,3-Trimethylbenzene	-0.16
Mono-Aromatics	1-Methyl-2-ethylbenzene	-0.16
Di-Olefins	2-Methyl-1,3-Butadiene	-0.13
Mono-Aromatics	1,3,5-Trimethylbenzene	-0.12
Mono-Aromatics	1-Methyl-4-ethylbenzene	-0.11
Mono-Aromatics	n-Propylbenzene	-0.09
Di-Olefins	1c/t,4-Hexadiene	-0.07
Di-Olefins	1t,3-Pentadiene	-0.06

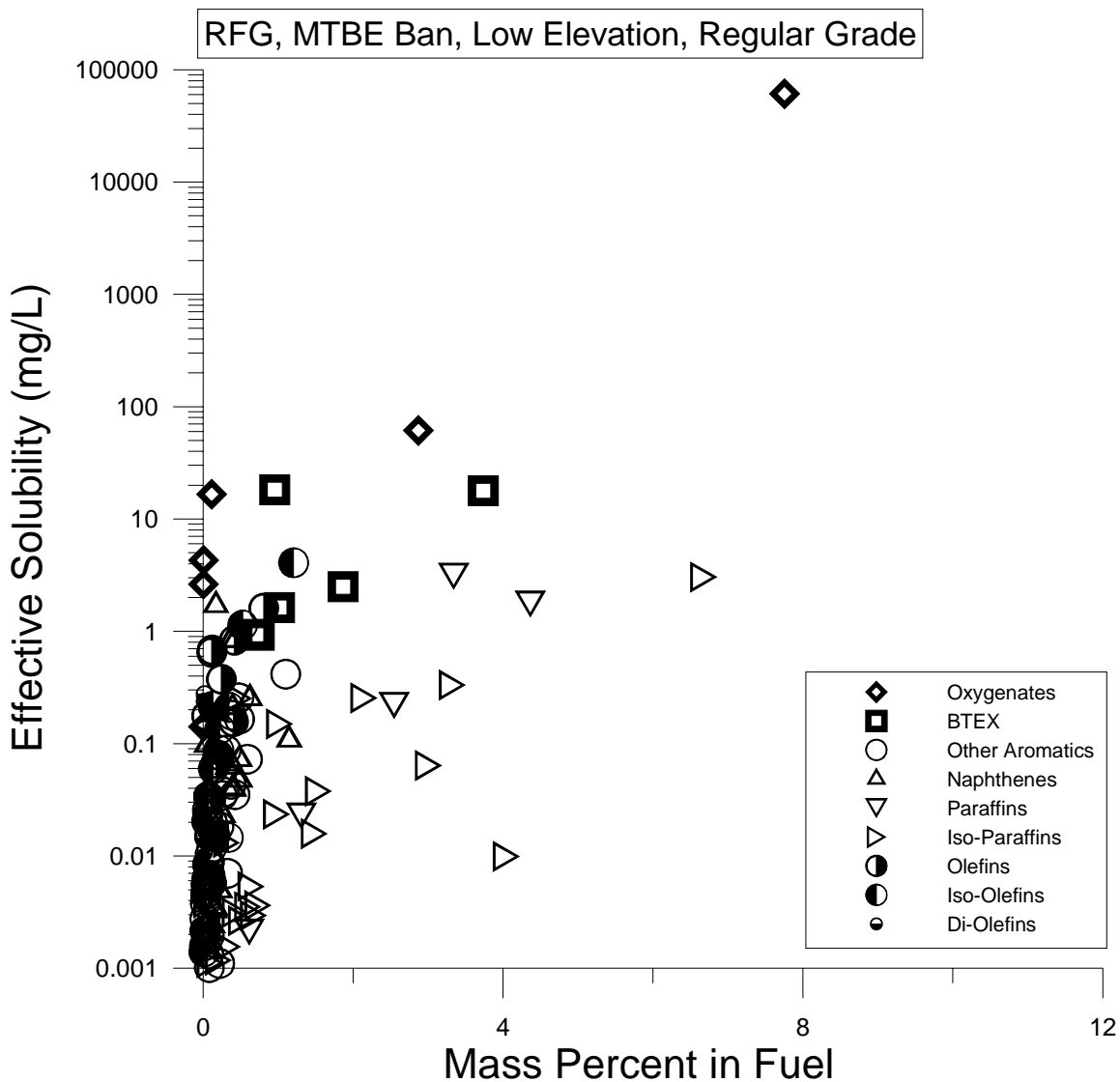


Figure 28 Effective Solubility for hydrocarbon groups and oxygenates for the Illinois regular gasolines (low elevation, MTBE ban, RFG).

Table 17 Predicted *higher* aqueous constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non-MTBE ban RFG gasoline.

Class	Chemical	Change in Concentration (mg/L)
Oxygenates	Ethanol	61080
Oxygenates	i-Butanol	2.69
Mono-Aromatics	Benzene	2.57
Paraffin	n-Butane	1.47
Iso-Paraffins	i-Pentane	0.60
Paraffin	n-Pentane	0.39

Table 18 Predicted *lower* aqueous constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non-MTBE ban RFG gasoline.

Class	Chemical	Change in Concentration (mg/L)
Oxygenates	Methyl-t-butylether	-1606
Oxygenates	t-Butanol	-44.28
Mono-Aromatics	Toluene	-10.26
Oxygenates	TAME	-5.36
Oxygenates	2-Butanol	-4.70
Oxygenates	n-Butanol	-3.94
Oxygenates	i-Propanol	-3.78
Oxygenates	n-Propanol	-2.92
Mono-Aromatics	m-Xylene	-1.73
Mono-Aromatics	o-Xylene	-1.16
Mono-Aromatics	Ethylbenzene	-1.05
Mono-Aromatics	p-Xylene	-0.73
Naphtheno-Olefins	Cyclopentene	-0.45
n-Olefins	c-Butene-2	-0.44
n-Olefins	t-Butene-2	-0.41
n-Olefins	t-Pentene-2	-0.33
Iso-Olefins	2-Methylbutene-2	-0.32
Naphtheno-Olefins	1-Methylcyclopentene	-0.28
Mono-Aromatics	1,2,4-Trimethylbenzene	-0.26
Mono-Aromatics	1-Methyl-3-ethylbenzene	-0.20
n-Olefins	c-Pentene-2	-0.17
Mono-Naphthenes	Cyclopentane	-0.16
n-Olefins	Pentene-1	-0.15
Iso-Olefins	2-Methylbutene-1	-0.14
Mono-Aromatics	1-Methyl-2-ethylbenzene	-0.11
Indenes	2,3-Dihydroindene	-0.10
Mono-Aromatics	1,2,3-Trimethylbenzene	-0.09
Mono-Aromatics	1-Methyl-4-ethylbenzene	-0.07
Di-Olefins	2-Methyl-1,3-Butadiene	-0.07
Mono-Aromatics	1,3,5-Trimethylbenzene	-0.07
Iso-Paraffins	3-Methylpentane	-0.07
Iso-Paraffins	2-Methylpentane	-0.06
Iso-Olefins	2-Methylpentene-2	-0.06
n-Olefins	t-Hexene-2	-0.06
Di-Olefins	1t,3-Pentadiene	-0.06

Predicted Gas Phase Concentrations

The predicted gas phase concentrations resulting from direct contact between the fuel and air for the average conventional, no-MTBE-ban, low-elevation, regular-grade gasoline are shown in Figure 29. The highest concentrations (greater than 100 mg/L) were found from paraffins and iso-paraffins. In the next lower category (10 mg/L to 100 mg/L) are more paraffins, n-olefins, and iso-olefins. BTEX and other classes of compounds appear between 1 mg/L and 10 mg/L.

Figure 30 shows the average gas phase concentrations from the similar RFG, no-MTBE-ban samples. The main, obvious difference in the plots is the presence of MTBE at high weight % and high concentration. Tables 19, 20, and 21 show the increased and decreased gas phase concentrations for a shift from conventional to RFG fuel. MTBE, and a variety of iso-olefins have increased concentration, while various aromatics and others decrease.

Concentrations for Illinois samples (MTBE ban) are shown in Figure 31. The general pattern shown is similar to the other fuels. Ethanol has a lowered concentration compared to MTBE. Concentrations that increase (Table 22) are ethanol and several iso-paraffins, indicating a need to increase the octane number, among others. Decreasing (Tables 23 and 24) concentrations were noted for MTBE and a variety of other compounds.

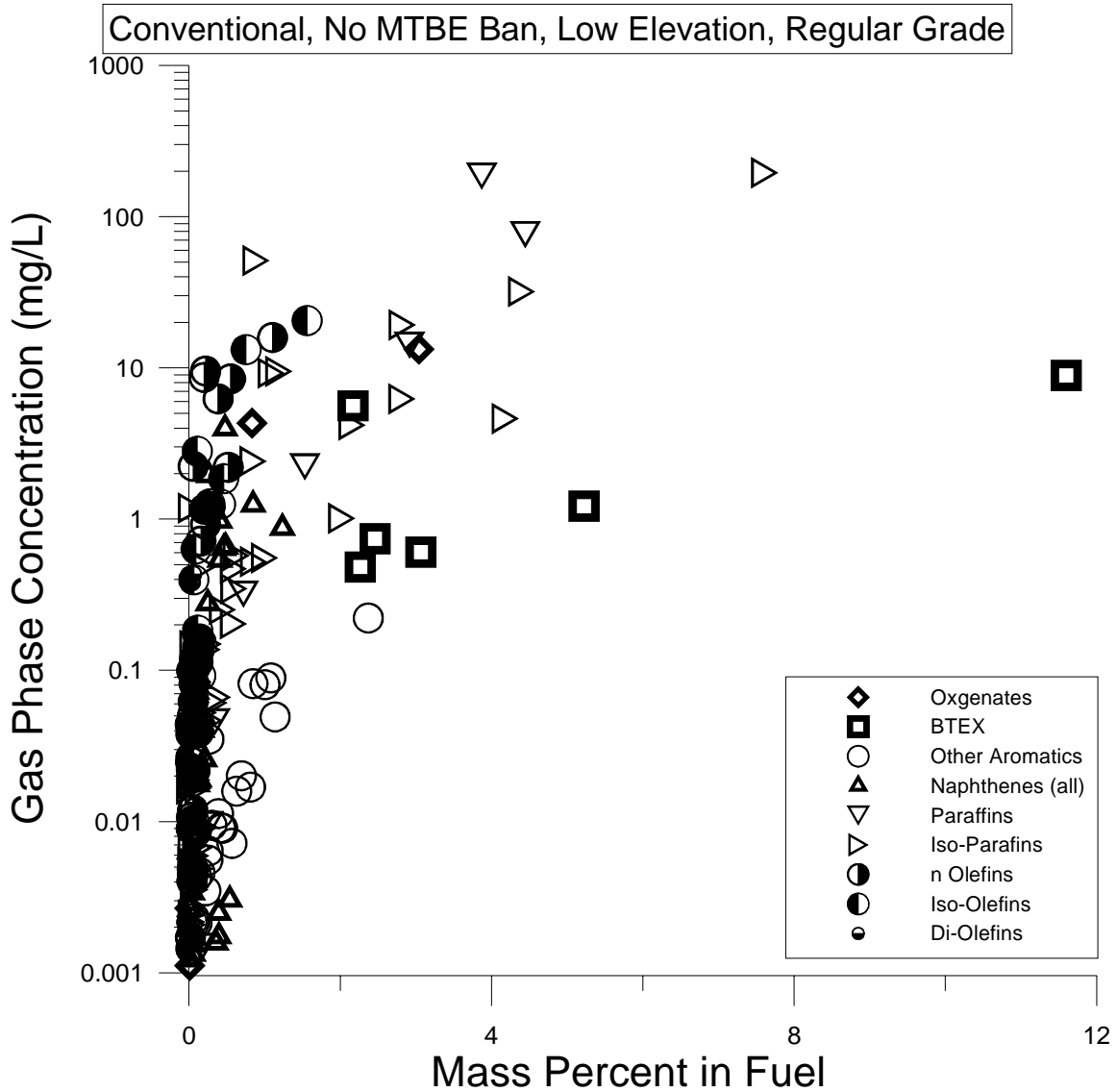


Figure 29 Gas phase concentration of hydrocarbon groups and oxygenates for the average conventional, low-elevation, no-MTBE-ban regular gasolines.

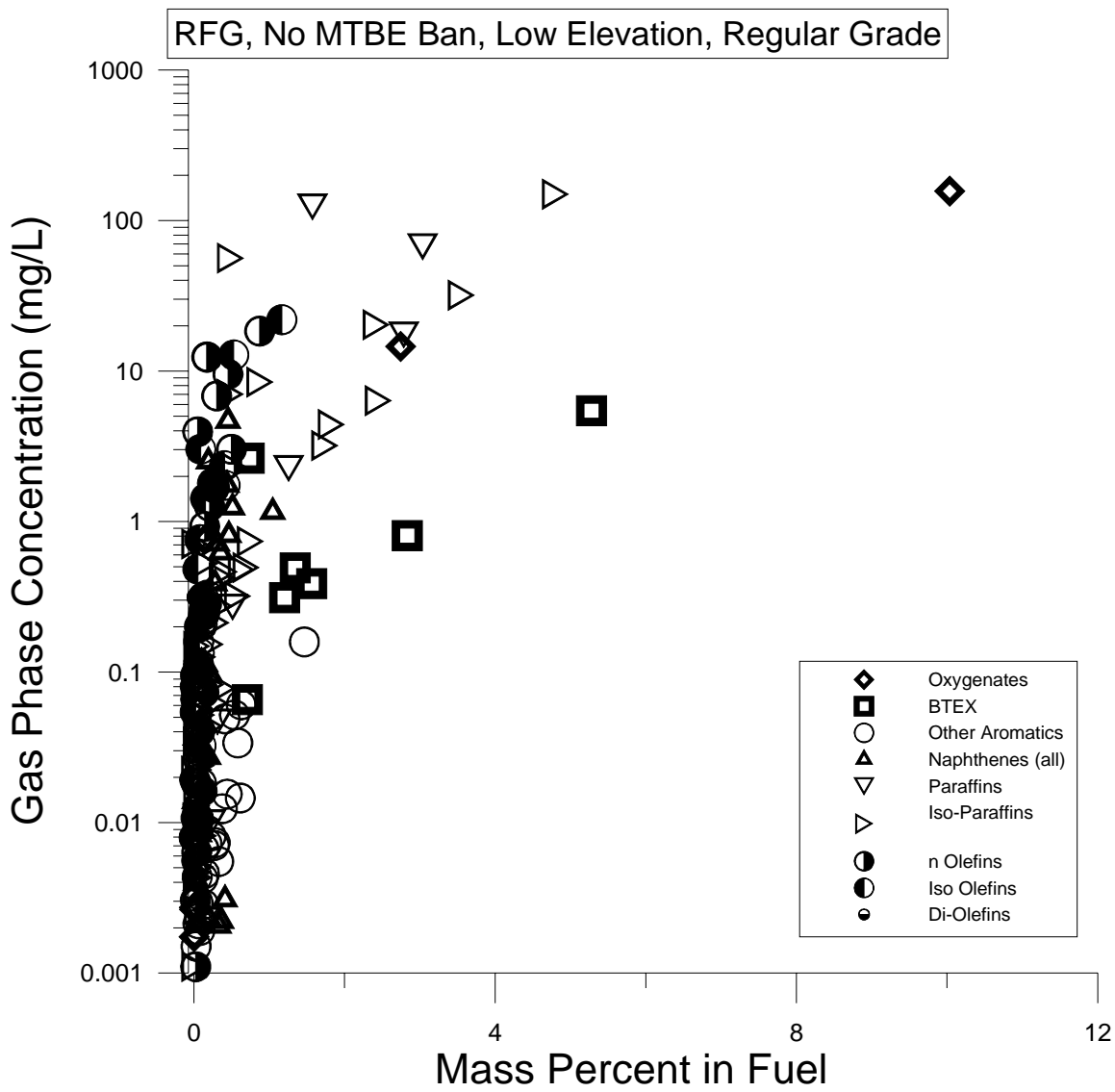


Figure 30 Gas phase concentration of hydrocarbon groups and oxygenates for the average reformulated, low-elevation, no-MTBE-ban regular gasolines.

Table 19 Predicted *higher* gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.

Class	Chemical	Change in Concentration (mg/L)
Oxygenates	Methyl-t-butylether	152.5
Iso-Paraffins	i-Butane	4.97
n-Olefins	t-Butene-2	3.72
Paraffin	n-Hexane	3.18
n-Olefins	c-Butene-2	2.82
n-Olefins	t-Pentene-2	2.50
n-Olefins	Butene-1	1.72
Iso-Olefins	2-Methylbutene-2	1.37
n-Olefins	c-Pentene-2	1.06
Iso-Paraffins	3-Methylpentane	1.03
n-Olefins	t-Hexene-2	0.82
Naphtheno-Olefins	1-Methylcyclopentene	0.78
Mono-Naphthenes	Cyclopentane	0.68
Oxygenates	TAME	0.67
Iso-Olefins	2-Methylpentene-1	0.60
n-Olefins	Pentene-1	0.58
Naphtheno-Olefins	Cyclopentene	0.55
Iso-Olefins	2-Methylpentene-2	0.50
n-Olefins	t-Hexene-3	0.50
Iso-Olefins	3,3-Dimethylpentene-1	0.48
Iso-Olefins	3-Methyl-c-pentene-2	0.41
n-Olefins	c-Hexene-2	0.34
Mono-Naphthenes	Methylcyclohexane	0.28
Iso-Olefins	4-Methyl-t-pentene-2	0.26
Iso-Olefins	2-Methyl-2-hexene	0.24
Iso-Olefins	3-Methyl-t-hexene-3	0.23
Iso-Paraffins	3-Methylhexane	0.22
Iso-Olefins	3-Methylbutene-1	0.21
n-Olefins	Hexene-1	0.21
Mono-Naphthenes	1c,3-Dimethylcyclopentane	0.15
Iso-Olefins	5-Methyl-c-hexene-2	0.13
n-Olefins	c-Heptene-2	0.13
Iso-Olefins	4-Methyl-t/c-hexene-2	0.13
Iso-Olefins	4-Methylpentene-1	0.12
n-Olefins	t-Heptene-3	0.12
n-Olefins	c-Heptene-3	0.12
Iso-Paraffins	2,2,5-Trimethylhexane	0.12
n-Olefins	t-Heptene-2	0.11
Mono-Naphthenes	Ethylcyclopentane	0.11
Iso-Paraffins	2-Methylhexane	0.11
Iso-Olefins	3-Methyl-t-hexene-2	0.11
Iso-Olefins	2-Methyl-t-hexene-3	0.10
Iso-Olefins	3-Methyl-c-hexene-2	0.10
Iso-Olefins	1,5-DM-Cyclopentene	0.09

Table 20 (Continuation) Predicted *higher* gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.

Class	Chemical	Change in Concentration (mg/L)
Naphtheno-Olefins	Cyclohexene	0.09
Iso-Olefins	4-Methyl-c-pentene-2	0.09
Mono-Naphthenes	1t,3-Dimethylcyclopentane	0.08
Iso-Olefins	2-Ethyl-3-methylbutene-1	0.08
Mono-Naphthenes	1t,2-Dimethylcyclopentane	0.07
Mono-Naphthenes	1,1,2-Trimethylcyclopentane	0.06
Oxygenates	t-Butanol	0.05
Iso-Olefins	3-Methyl-c-hexene-3	0.05
Iso-Olefins	3-Ethylpentene-2	0.05

Table 21 Predicted *lower* gas phase constituent concentrations (>0.05 mg/L) of RFG (regular grade, low elevation) in comparison to conventional gasoline.

Class	Chemical	Change in Concentration (mg/L)
Paraffin	n-Butane	-62.91
Iso-Paraffins	i-Pentane	-45.31
Paraffin	n-Pentane	-9.36
Mono-Aromatics	Toluene	-3.46
Mono-Aromatics	Benzene	-2.98
Iso-Paraffins	2,2-Dimethylbutane	-2.22
Iso-Paraffins	2,2,4-Trimethylpentane	-1.43
Iso-Paraffins	2,3-Dimethylbutane	-1.05
Iso-Paraffins	2,2-Dimethylpropane	-0.47
Iso-Olefins	2-Methylbutene-1	-0.43
Mono-Aromatics	m-Xylene	-0.41
Iso-Paraffins	2,3,4-Trimethylpentane	-0.27
Mono-Aromatics	Ethylbenzene	-0.25
Mono-Aromatics	o-Xylene	-0.22
Mono-Aromatics	p-Xylene	-0.17
Di-Olefins	1c/t,4-Hexadiene	-0.16
Iso-paraffins	2,4-Dimethylpentane	-0.14
Di-Olefins	2-Methyl-1,3-Butadiene	-0.12
Iso-paraffins	2,4-Dimethylhexane	-0.11
Mono-Aromatics	1,2,4-Trimethylbenzene	-0.07
Iso-paraffins	2,3-Dimethylhexane	-0.07
Di-Olefins	1t,3-Pentadiene	-0.07
Mono-Aromatics	1-Methyl-3-ethylbenzene	-0.06
Iso-paraffins	3-Methylheptane	-0.06
Iso-paraffins	2-Methylpentane	-0.06
Paraffin	n-Octane	-0.05

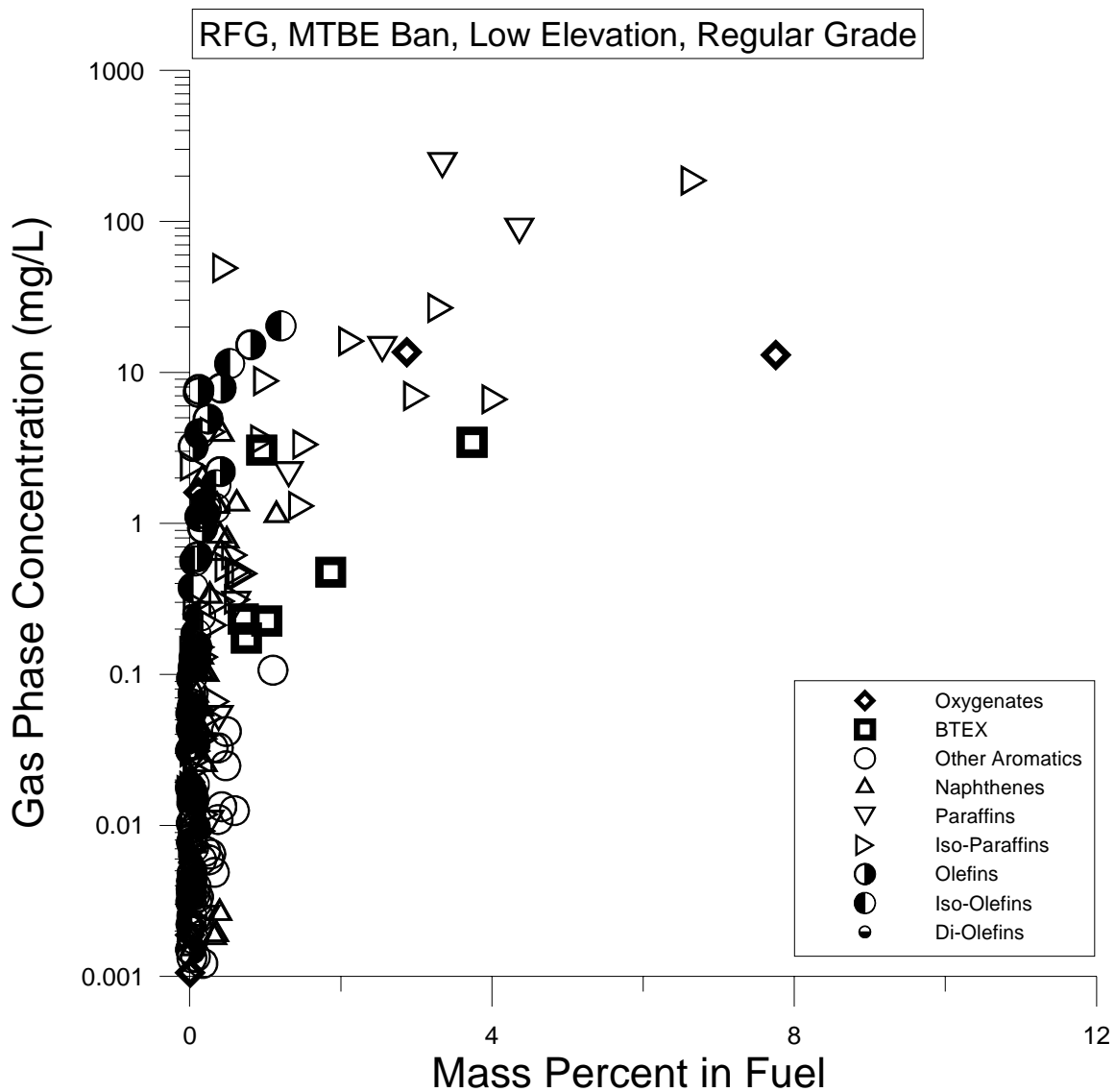


Figure 31 Gas phase concentration for hydrocarbon groups and oxygenates for the Illinois regular gasolines (low elevation, MTBE ban, RFG).

Table 22 Predicted *higher* gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.

Class	Chemical	Change in Concentration (mg/L)
Paraffin	n-Butane	111.13
Iso-paraffins	i-Pentane	36.77
Paraffin	n-Pentane	19.22
Oxygenates	Ethanol	13.04
Iso-paraffins	2,2,4-Trimethylpentane	3.45
Iso-paraffins	2,2-Dimethylpropane	1.68
Iso-paraffins	2,4-Dimethylpentane	1.33
Iso-Olefins	3-Methylbutene-1	0.90
Iso-paraffins	2-Methylhexane	0.62
Iso-paraffins	2,3,4-Trimethylpentane	0.57
Iso-paraffins	2-Methyl-3-ethylpentane	0.48
Mono-Aromatics	Benzene	0.43
Iso-paraffins	2,3-Dimethylbutane	0.34
Iso-paraffins	2,4-Dimethylhexane	0.15
Mono-Naphthenes	1c,2c,3-Trimethylcyclopentane	0.13
Mono-Naphthenes	1t,2-Dimethylcyclopentane	0.12
Mono-Naphthenes	Cyclohexane	0.11
Mono-Naphthenes	1t,4-Dimethylcyclohexane	0.10
Iso-paraffins	2,5-Dimethylhexane	0.09
Mono-Naphthenes	3c-Ethylmethylcyclopentane	0.06
Iso-Olefins	2-Methyl-c-hexene-3	0.06

Table 23 Predicted *lower* gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.

Class	Chemical	Change in Concentration (mg/L)
Oxygenates	Methyl-t-butylether	-155.16
Iso-paraffins	i-Butane	-7.11
Iso-paraffins	2-Methylpentane	-5.17
n-Olefins	c-Butene-2	-5.00
n-Olefins	t-Butene-2	-4.66
Iso-paraffins	3-Methylpentane	-4.20
n-Olefins	t-Pentene-2	-3.15
Paraffin	n-Hexane	-3.14
Iso-paraffins	2,2-Dimethylbutane	-2.99
Mono-Aromatics	Toluene	-1.99
n-Olefins	Pentene-1	-1.92
n-Olefins	c-Pentene-2	-1.65
Iso-Olefins	2-Methylbutene-2	-1.62
Iso-Olefins	2-Methylbutene-1	-1.36
Iso-paraffins	3-Methylhexane	-1.08
n-Olefins	t-Hexene-2	-0.82
Mono-Naphthenes	Cyclopentane	-0.75
Oxygenates	TAME	-0.74
n-Olefins	Butene-1	-0.72
n-Olefins	t-Hexene-3	-0.56
Iso-Olefins	2-Methylpentene-2	-0.54
Naphtheno-Olefins	Cyclopentene	-0.52
Iso-Olefins	3,3-Dimethylpentene-1	-0.48
Naphtheno-Olefins	1-Methylcyclopentene	-0.45
Iso-Olefins	3-Methyl-c-pentene-2	-0.45
Iso-Olefins	2-Methylpentene-1	-0.44
n-Olefins	c-Hexene-2	-0.33
Mono-Aromatics	m-Xylene	-0.33
n-Olefins	Hexene-1	-0.33
Iso-Olefins	4-Methyl-t-pentene-2	-0.31
Mono-Aromatics	Ethylbenzene	-0.26
Iso-paraffins	3,3-Dimethylpentane	-0.24
Iso-Olefins	2-Methyl-2-hexene	-0.22
Iso-Olefins	4-Methylpentene-1	-0.20
Mono-Aromatics	o-Xylene	-0.16
Iso-Olefins	5-Methyl-c-hexene-2	-0.16
Paraffin	n-Heptane	-0.15
Mono-Aromatics	p-Xylene	-0.14
n-Olefins	c-Heptene-3	-0.13
Iso-Olefins	3-Methyl-t-hexene-2	-0.13
n-Olefins	c-Heptene-2	-0.13
Iso-Olefins	4-Methyl-t/c-hexene-2	-0.13
n-Olefins	t-Heptene-2	-0.12
Iso-Olefins	3-Methyl-c-hexene-2	-0.11
Iso-Olefins	4-Methyl-c-pentene-2	-0.11

Table 24 (Continuation) Predicted *lower* gas phase constituent concentrations (>0.05 mg/L) of MTBE-ban RFG (regular grade, low elevation) in comparison to non MTBE-ban RFG.

Class	Chemical	Change in Concentration (mg/L)
n-Olefins	t-Heptene-3	-0.10
Iso-Olefins	2-Methyl-t-hexene-3	-0.09
Iso-Olefins	1,5-DM-Cyclopentene	-0.09
Iso-Olefins	3-Methyl-t-hexene-3	-0.08
Iso-Olefins	2-Methyl-c-hexene-3	-0.08
Iso-Olefins	2-Ethyl-3-methylbutene-1	-0.07
Mono-Naphthenes	1,1,2-Trimethylcyclopentane	-0.07
Di-Olefins	2-Methyl-1,3-Butadiene	-0.06
Di-Olefins	1t,3-Pentadiene	-0.06
Naphtheno-Olefins	Cyclohexene	-0.06
Iso-Olefins	3-Methyl-c-hexene-3	-0.06
Oxygenates	t-Butanol	-0.06
Mono-Naphthenes	Ethylcyclopentane	-0.06
Iso-Olefins	3-Ethylpentene-2	-0.05
Mono-Aromatics	1-Methyl-3-ethylbenzene	-0.05

Figure 32 shows one final example where the predicted gas phase concentration of the average conventional, no MTBE ban, low elevation, regular grade gasoline components were plotted against their effective solubilities. For reference a line of equal gas phase concentration and effective solubility was plotted. The gas phase results are based upon vapor pressure (i.e., direct volatilization from the fuel) and the oxygenates plotted below the equal-concentration line, as did many naphthenes and aromatics. The BTEX compounds, paraffins, Iso-Paraffins and all three types of olefins plotted above the line, indicating that their presence is greater in the gas phase than in the water phase. Notably, in the preceding figures of gas phase concentration (Figures 29, 30, and 31) there was a shift upward of paraffins and Iso-Paraffins relative to the plots for the effective solubility (Figures 26, 27 and 28).

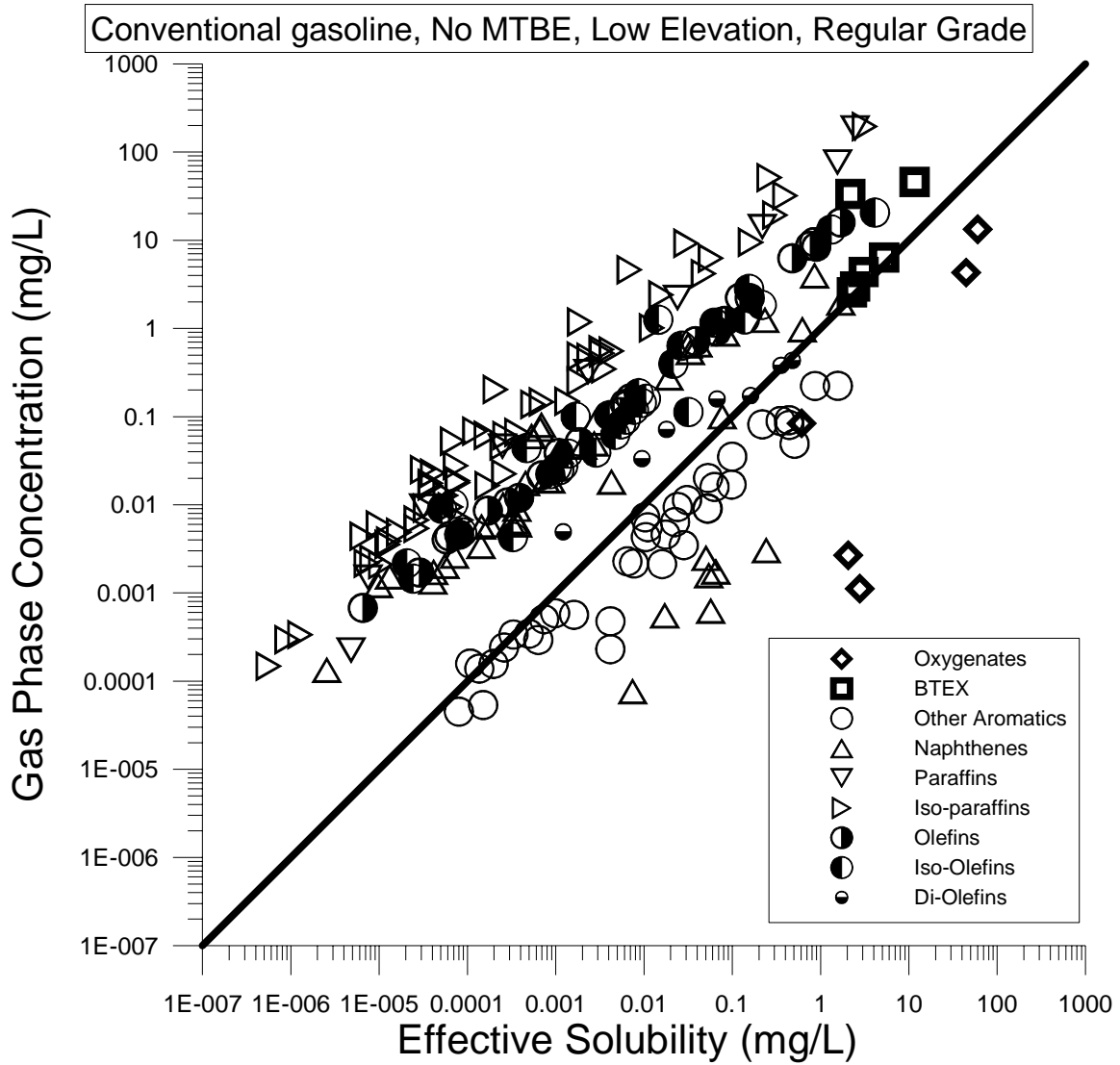


Figure 32 Scatter plot of predicted gas phase concentration versus the effective solubility of the average conventional, low-elevation, no-MTBE-ban, regular gasoline.

Discussion and Conclusions

Data were collected on 312 components of gasoline based on 96 samples collected from across the United States. The sample locations were selected to include RFG and conventional gasoline, high and low elevation, regular and premium grades, likely ethanol-using locations and MTBE-ban states. Evaluation of single compound concentrations lead to several findings:

1) The benzene concentration of Federal RFG fuels were consistently below 1% by weight. This result follows from federal requirements and agrees with data collected by EPA for compliance (US EPA, 2005). In conventional fuels, higher benzene concentrations of roughly 0.5% to 3% were found. These values generally agree with data generated by industry (Dickson, 2004a, 2004b). Data on conventional fuels available from EPA show the average benzene concentration to range from 1.11 % to 1.17% by volume for summer and 1.08% to 1.15% by volume for winter for the years from 1997 to 2002. These data are not necessarily in conflict with the current study data as the averages are based on many more samples and include more locations across the country. The EPA data show that benzene ranged from 0.01 % to about 5.0% by volume for summer and winter samples in 2002 and 2003 (John Weihrauch, U.S. EPA, Office of Transportation and Air Quality, 2005, personal communication).

A) In Colorado the state oxygenate requirement did not limit the amount of benzene in the fuel while it required an oxygen content of 3.1 %, which is higher than the federal requirement.

2) Where MTBE bans are in place, ethanol has replaced MTBE as the main oxygenate in use. Ethanol was not found in other samples suggesting it was used only where needed to meet an oxygenate requirement.

3) Only low concentrations of TBA were found in the fuels and were associated with MTBE.

The gasoline data generated for this study represent seven combinations of elevation, RFG requirements, and MTBE ban status. The regular gasolines showed more similarity given a particular combination of these factors, than did the premiums. This result suggests that refiners use more approaches to raise the octane rating of the fuels. Statistical analysis of the data support the division of the samples by grade, elevation, MTBE ban status and RFG requirements for some combinations of the factors. Differences due to state RFG requirements, regional use of oxygenates for octane enhancement and lack of commonality among high elevation fuels, require specific local data to achieve the goal of prediction of release composition. The need to predict release composition follows from the lack of detailed characterization of most fuels and lag time between release and field sampling. Both of these are expected to continue in the future.

The environmental impact of these fuels depends on the exposure pathway. When contact with contaminated soils is a significant pathway, the components of the fuel are themselves directly important. If exposure occurs through ground water or soil gas, then the concentrations in these media indicate the importance of the various components of the fuel. Effective

solubilities and gas phase concentrations were approximated from estimates of solubility vapor pressure and the individual gasoline component concentration data. In several examples, these showed that a smaller number of components result in high (say greater than 1 mg/L) concentrations in these media. For ground water the highest concentrations were found for the oxygenates and BTEX. For the gas phase concentration, the highest were found for paraffins and oxygenates, with the BTEX occurring at lower concentration. These only provide a rough guide, because the toxicity of the compound and the exposure ultimately determines the significance of these chemicals.

References

- 40 CFR 80, Code of Federal Regulations, Title 40 Protection of Environment, Part 80 – Regulations of Fuels and Fuel Additives
- American Petroleum Institute, 2004, State MTBE Activity Laws, 6/14/04, Unpublished document.
- American Society for Testing and Materials, 1958, Knocking Characteristics of Pure Hydrocarbons, ASTM Special Publication No. 225, Philadelphia, 96 pp.
- American Society for Testing and Materials, 2002, Standard Test Method for Determination of Individual Components in Spark Ignition Engine Fuels by 100 Meter Capillary High Resolution Gas Chromatography, ASTM International, West Conshohocken, Pennsylvania, D 6729-1.
- California Air Resources Board, 2003, The California Reformulated Gasoline Regulations, Title 13, California Code of Regulations, Sections 2250-2273.5, May 1, 2003.
- Canadian Petroleum Products Institute, 1994, Composition of Canadian Summer and Winter Gasolines 1993, CPPI Report No. 94-5.
- Chevron U.S.A. Inc, 2005, Motor Gasolines Technical Review, www.chevron.com.
- Colorado, 2002, Regulation 13, The Reduction of Carbon Monoxide Emissions from Gasoline Powered Motor Vehicles through the use of Oxygenated Gasolines.
- Dickson, C. L., 2004, Motor Gasolines, Winter 2003-04, Northrup Grumman Mission Systems, Bartlesville, Oklahoma, NGMS-235 PPS 2004/3, 67pp.
- Dickson, C. L., 2004, Motor Gasolines, Summer 2003, Northrup Grumman Mission Systems, Bartlesville, Oklahoma, NGMS-233 PPS 2004/1, 67pp.
- Gustafson, J.B., J.G. Tell, and D. Orem, 1997, Selection of Representative TPH Fractions Based on Fate and Transport Considerations, Total Petroleum Hydrocarbons Criteria Working Group Series Volume 3, Amherst Scientific Publishers, Amherst Massachusetts, 102pp.
- Hilal, S.H., S.W. Karickhoff and L.A. Carreira 2003, Verification and Validation of the SPARC Model, US EPA, Washington, D.C, EPA/600/R-03/033.
- Jokuty, P., S. Whiticar, Z. Wang, M. Fingas, B. Fieldhouse, P. Lambert and J. Mullin, Properties of Crude Oils and Oil Products, Manuscript Report EE-165, Environmental Protection Service, Environment Canada, Ottawa, Ontario, 1999. Available at <http://www.etc-cte.ec.gc.ca/databases/OilProperties>

- Jolliffe, I.T. 2002. Principle component analysis. Springer-Verlag, New York. 487 p.
- Jolliffe, I.T. 1970. Redundant variables in multivariate analysis. Unpublished PhD. Dissertation. University of Sussex.
- Jolliffe, I.T. 1972. Discarding variables in principal component analysis 1: Artificial data. *Appl. Statist* (21):160-173 p.
- Jolliffe, I.T. 1972. Discarding variables in principal component analysis 2: Real data. *Appl. Statist* (22):21-31 p.
- Karickhoff, S.W., V.K. McDaniel, C. Melton, A. N. Vellino, D. E. Nute, L.A. Carriera, 1991, Predicting chemical reactivity by computer, *Environmental Toxicology and Chemistry*, 10, 1405-1416.
- Kendall, M. 1975. Multivariate analysis. Hafner Press, New York. 210 p.
- Khattree, R., and D. N. Naik. 2000. Multivariate data reduction and discrimination with sas software. SAS Institute, Cary, NC. 558 p.
- Maine Department of Environmental Protection, 2005, 2004 Maine Fuels Report, Augusta, Maine.
- Meusinger, R and R. Moros, 2001, Determination of octane numbers of gasoline compounds from their chemical structure by ^{13}C NMR spectroscopy and neural networks, *Fuel*, 80, 613-621.
- Rencher, A.C., 1995, *Methods of Multivariate Analysis*, Wiley, New York, 627pp.
- Schwarzenbach, R.P, P.M. Gschwend, and D.M. Imboden, 2003, *Environmental Organic Chemistry*, 2nd ed., Wiley, New York, 1313pp.
- Shaw, D.G, 1989, *Solubility Data Series, Hydrocarbons with Water and Sea Water, Part 1: Hydrocarbons C₅ to C₇*, Volume 37, International Union of Pure and Applied Chemistry, Pergammon Press.
- Shaw, D.G, 1989, *Solubility Data Series, Hydrocarbons with Water and Sea Water, Part 2: Hydrocarbons C₈ to C₃₆*, Volume 38, International Union of Pure and Applied Chemistry, Pergammon Press.
- Stout, S.A., A.D. Uhler, R.M. Uhler, E.M. Healey, and K.J. McCarthy, 2003, Detailed chemical fingerprinting of gasoline for environmental forensic investigations: Part 3. Applications to environmental forensics investigations, *Soil, Sediment and Water*, March/April, 16-18.
- Strickers, D.E., 2002, Octane and the environment, *The Science of the Total Environment*, 299,

37-56.

United States Environmental Protection Agency, 2005, Office of Transportation and Air Quality, <http://www.epa.gov/otaq>.

United States Environmental Protection Agency, 2005b, Where You Live Map of Current RFG Areas, Office of Transportation and Air Quality.

Washington State, 1996, Washington Administrative Code: Chapter 173-492 WAC: Motor Fuel Specifications for Oxygenated Gasoline.

Weaver, J.W., R.J. Charbeneau, J.D. Tauxe, B.K. Lien, and J.B. Provost, 1994, The Hydrocarbon Spill Screening Model (HSSM) Volume 1: User's Guide, US EPA, EPA/600/R-94/039a.

Weaver, James W., 2004, On-line Tools for Assessing Petroleum Releases, United States Environmental Protection Agency, Washington, D.C., EPA 600/R-04/101.

Appendix: Theoretical Basis of the Principal Components Analysis

Following the theory presented by Jolliffe (2002), let X be a vector of p random variables, such that the variances of the p random variables and the structure of the covariances or correlations between the p variables are of interest. Unless p is small, it is often unbeneficial to look at the p variances and all of the $0.5p(p-1)$ correlations or covariances. An alternative approach is to find a few (less than p) variables that preserve most of the information given by these variances and correlations or covariances. The initial step is to find a linear function $\alpha_1^T X$ of the elements of X having maximum variance, where α_1 is a vector of p constants $\alpha_{11}, \alpha_{12}, \dots, \alpha_{1p}$, and T denotes the transpose, such that

$$\alpha_1^T X = \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1p}x_p = \sum_j^p \alpha_{1j}x_j \quad (5)$$

Next, find a linear function $\alpha_s^T X$ which is uncorrelated with $\alpha_1^T X$ having maximum variance, and continuing repeating this sequence so that at the k th stage a linear function $\alpha_k^T X$ is found that has maximum variance subject to being uncorrelated with $\alpha_1^T X, \alpha_s^T X, \dots, \alpha_k^T X$. Up to p PC could be found, however it is hoped that the majority of the variation in X will be accounted for by m PC, where m is less than p .

The first step in PCA is deciding whether to use the correlation matrix or covariance matrix. Use of the correlation matrix forces all of the variables to have equal variance and this in turn may defeat the purpose of identifying those variables that contribute more significantly to the total variability (Khattree and Naik, 2000). The PC of a PCA based on covariance matrices are also insensitive to units of measurement on the same scale. Since all of the gasoline variables are measured on a percent weight basis, ensures the PC will be interpretable and unbiased by those variables having large variance. PCA in this work was done utilizing the covariance matrix of the variables.

There are two commonly used methods for determining the number of PC to select. The first, and most commonly used, is based on the cumulative proportion of total variance. An appropriate minimum percentage of total variation desired to be explained by the PCA is prespecified, and the smallest number of principal components that satisfies this criterion is selected. The prespecified percentage is usually taken to be 95 percent variation explained. The second method is graphical and uses what is called a scree diagram. A scree diagram is a plot of the variances against the p PC. From the plot, the number of m PC to be selected is determined in such a way that the slope of the graph is steep to the left of m , but at the same time not steep to the right. The idea being that the number of PC to be selected is such that the differences between consecutive variances are becoming increasingly smaller. Both of the techniques described above will be used in determining an appropriate number of PC.

When the number of variables p is large, it is often the case that a subset of m variables, where $m < p$ contains virtually all the information available in all p variables. It is useful then to select an appropriate value of m , and then determine which subset of m variables are best. It has

been cited that if X can successfully be described by m PC, then X can be replaced by a subset of m variables (Jolliffe 1970, 1972, 1973). The variable corresponding to the maximum of the absolute values of the coefficients in the first PC is selected first. The second variable selected is that variable (if not already selected) whose coefficient in the second PC has the maximum absolute value. If the variable corresponding to the maximum coefficient was already selected then we choose the variable corresponding to the next maximum. Proceeding this way, we select a subset of required size from the list of original variables.

Given the data have p variables the observations can be plotted as points in p -dimensional space. Two-dimensional plots are of use in detecting patterns in the data. If the data lie close to a two dimensional subspace, plots of the $m \leq p$ PC will provide an adequate representation for detecting patterns or clusters of data since the majority of the variation is accounted for by the m PC. Graphical analysis will be conducted to determine if unique trends among the samples are observed. For these data, we looked for patterns among location, octane number, MTBE ban (Yes or No), RFG or conventional gasoline, brand, and altitude.