ANALYSIS OF PHOTOCHEMICAL ASSESSMENT MONITORING STATION (PAMS) DATA TO EVALUATE A REFORMULATED GASOLINE (RFG) EFFECT

FINAL REPORT
STI-997350-1774-FR2

Prepared for:

U.S. Environmental Protection Agency
Office of Mobile Sources
Washington, DC

By:

Hilary H. Main
Paul T. Roberts
Richard Reiss
Sonoma Technology, Inc.
Santa Rosa, CA

April 1998

DISCLAIMER

The statements and conclusions in this report are not necessarily those of the Office of Mobile Sources, the Environmental Protection Agency, or their employees. The mention of commercial products, their source, or their use in connection with material reported herein is not to be construed as actual or implied endorsement of such products.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

TABLE OF CONTENTS (Concluded)

LIST OF FIGURES

LIST OF FIGURES (Concluded)

LIST OF TABLES

LIST OF TABLES (Concluded)

EXECUTIVE SUMMARY

The reformulation of motor vehicle fuel began in late 1992 in California and in early 1995 outside California. Under the U.S. Environmental Protection Agency (EPA) Office of Mobile Sources (OMS) sponsorship, this study evaluated the use of Photochemical Assessment Monitoring Station (PAMS) data to provide some early confirmation of reformulated gasoline's (RFG) ambient air quality benefits in a number of regions around the country. The objective of this analysis was to determine if the PAMS measurements indicate real reductions in ambient concentrations of ozone precursors and toxics, rather than fluctuations, and to what extent these reductions in ambient concentrations can be linked to the implementation of the RFG program. Results of this investigation indicate that RFG has reduced some ambient hydrocarbon concentrations.

ES.1 BACKGROUND

In accordance with the Clean Air Act Amendments of 1990, the use of RFG was required in the nine worst ozone nonattainment areas of the country. Other areas of the country also opted to implement RFG programs to supplement their clean air plans. In 1991, the National Academy of Sciences (NAS) released a report entitled, "Rethinking the Ozone Problem in Urban and Regional Air Pollution," which criticized the EPA for failing to establish monitoring networks to adequately track trends in ozone precursor emissions, corroborate emission inventories, and support photochemical modeling. The PAMS program forms the basis of the EPA's response to recommendations from the NAS as well as the agency's response to the NAS report. Hydrocarbon and carbonyl data, as well as surface meteorological data, are collected at PAMS sites routinely. Many states implemented PAMS sites in 1994; similar data collection has been carried out in several areas for many years.

ES.2 REFORMULATED GASOLINE

This study focuses on the impact of the implementation of Phase I RFG on ambient concentrations. The principal differences between Phase I RFG and conventional gasoline include:

- Benzene levels are limited to no more than 1.0 percent by volume in RFG; in conventional gasoline, benzene can be as high as 5.0 percent by volume.

- A minimum 2.0 percent by weight of oxygen content is required in RFG. This requirement is typically met by either adding up to about 10 percent ethanol or 11 percent by volume methyl tertiary butyl ether (MTBE). Refiners are replacing aromatic hydrocarbons with high octane oxygenates to meet the oxygen content requirement and to lower aromatic hydrocarbon levels.

- The Reid Vapor Pressure (RVP) of the gasoline is reduced. Typically, butanes may be reduced and higher alkanes and aromatics may be added to decrease RVP.

- Required reductions of air toxics and VOC emissions by 15 percent result in RFG fuels having lower aromatic content compared to baseline fuels. Typically, RFG fuels contain 20 to 27 percent aromatics whereas baseline fuel aromatics content was 32 percent (summer) and 26 percent (winter) by volume.

Independent of fuel changes related to RFG, some areas of the country have modified gasoline's RVP and oxygen content. Oxygenate use is required in some areas of the country to help reduce wintertime carbon monoxide concentrations, and RVP regulations have been in place to help reduce hydrocarbon emissions that contribute to ozone formation. Therefore, a review of RFG's effectiveness needs to assess what best represents the "before and after" conditions for gasoline in the region to be studied (i.e., what was the gasoline's composition before and after RFG implementation). In addition, fuel characteristics vary from one region of the country to another, so fuel characteristics need to be specific to each area.

As part of this study, a significant effort was made to document fuel content before and after RFG implementation on a city-by-city basis. Three data sets proved useful for this investigation:

- American Automobile Manufacturers Association (AAMA, previously Motor Vehicle Manufacturers Association, MVMA) fuel survey results for 1994 and 1995.
- National Institute for Petroleum and Energy Research (NIPER) summer gasoline analyses for 1993 through 1996.
- RFG Survey Association fuel survey data for 1995 and 1996 collected at the retail level.

The AAMA fuel survey results for 1994 and 1995 are summarized in **Table ES-1** for six cities. Most of the metropolitan areas investigated in this study showed reductions in RVP, aromatic content, olefin content, and benzene with increases in the saturate (paraffin) content. The NIPER survey data for 1993 through 1996 were also investigated and qualitatively were consistent with the AAMA data. The compliance survey data showed that fuel composition continued to change after initial RFG implementation. For example, benzene levels in fuel showed some increases from 1995 to 1996 in the Boston, Washington, Milwaukee, and Chicago areas.

Table ES-1. Absolute change (volume percent) in average fuel composition between 1994 and 1995 (AAMA fuel survey results). Regular unleaded fuel.

City	RVP[a]	Aromatic	Olefins	Saturates	Benzene	MTBE
Boston, MA	-0.7	-2.2	-10	12	-1.0	8.6
Chicago, IL	-1.1	-0.9	-0.9	1.8	-0.5	[b]
Los Angeles, CA	-0.2	-3.1	-0.3	3.4	-0.6	10.6
New York, NY	-0.6	2.0	-7.9	6.0	-0.4	9.9
Philadelphia, PA	-0.5	-1.0	-2.8	3.8	-0.2	10.1
Washington, D.C.	-0.3	1.2	-0.4	-0.8	-0.1	10

[a] PSI
[b] Ethanol, rather than MTBE, is the primary oxygenate used in this area.

Fuel surveys and RFG implementation plans showed that most of the areas with PAMS sites had implemented RFG during early 1995 including Chicago, Boston, Milwaukee,

Philadelphia, New York City, Los Angeles, Providence, Washington, D.C., and Hartford, while Baton Rouge, and Harrington Beach, WI were two PAMS sites that were in areas that did not implement RFG.

ES.3 AMBIENT HYDROCARBON DATA

Thirteen PAMS data sets, encompassing areas with and without RFG implementation, were obtained, validated, and investigated. The following issues were considered in selecting data for downloading from AIRS (Aerometric Information Retrieval Systems, operated by EPA), during validation, and for use in analysis:

- Data Coverage. For the best assessment of changes in ambient species composition due to RFG implementation, data were required both before and after the change in fuel. Since most fuel changes occurred in early 1995, data for 1994 were imperative for this analysis. In some cases, 1993 data were also available and thus, data from two years before and two years after the change in fuel could be assessed. Data from both RFG and non-RFG sites were sought for comparison purposes.

- Measurement techniques. Different monitoring quality assurance procedures may significantly influence data availability and data quality. At two candidate sites, Bronx, NY and Chicago, IL, the measurement systems changed from 3-hr canisters in 1994 to 1-hr automated gas chromatography systems (auto-GCs) in 1995; this change in sampling could have influenced measured concentrations, and thus had an effect on the ability to discern a trend solely due to fuel changes.

- Quality of ambient data. It is critical to evaluate data quality and perform additional data validation, if necessary, before performing analyses, in order to ensure the analysis results are driven by robust data and not by invalid data. Note that some important species (e.g., MTBE and all alcohols) are not currently measured in the PAMS program.

All sites considered were PAMS Type II (maximum ozone precursor emissions impact) except Harrington Beach, WI which is a PAMS Type III site (downwind at maximum ozone concentration). Data from this site were used because of the paucity of data from other areas that did not implement RFG.

The issue of data validation was important for this project because of the use of ambient data in the analysis. Data validation is needed because serious errors in data analysis results can be caused by erroneous individual data values. Data validation and quality control is performed by the reporting agencies prior to submittal of the data to AIRS. However, our experience with these data has shown that additional inspection of data is imperative as many data problems have not been removed prior to data submittal. The objectives of the data validation process are to produce a database with values that are of a known quality; to evaluate the internal, spatial, temporal, and physical consistency of the data; and to intercompare data to identify errors, biases, or outliers. Outliers are defined as data values that are physically, spatially, or temporally inconsistent with the vast majority of observed values. The identification of outliers is particularly important and particularly difficult in the analysis of speciated hydrocarbon data, as the concentration distributions of these species are typically broad and some extreme values may be

the result of measurement error. Data validation is an ongoing process, and each new analysis provides an additional opportunity to gain confidence in the validity of the data. Problems were more prevalent in the early stages of the PAMS program (1993 and 1994) but improved during the 1995 and 1996 years.

ES.4 USEFUL INDICATOR SPECIES AND RATIOS

As a part of the Coordinating Research Council (CRC) Model Evaluation Feasibility Study (Stoeckenius et al., 1995), STI investigated toxic species, hydrocarbon, and carbonyl compound data collected in the South Coast Air Basin (SoCAB) of California from 1990 through 1993 (Main and Roberts, 1994). Based on predicted changes in the evaporative and exhaust emissions due to the introduction of California Phase I RFG (Ligocki and Yarwood, 1994), the temporal trends of selected species, species groups, and ratios were evaluated to assess the usefulness of these parameters as indicators of change in motor vehicle fuel composition (Main and Roberts, 1994; Main et al., 1995). Based on the SoCAB results, and our knowledge of the ambient PAMS data quality and availability, the current investigation focused on several species, species groups, and ratios (if available) as summarized in **Table ES-2**. Consensus among several indicators will give us more confidence in the study conclusions.

ES.4.1 CONCENTRATION AND WEIGHT FRACTION

There are several species and species group concentrations that may be good ambient indicators of fuel RFG changes. For example, since 15 percent reductions in smog-reducing emissions were required with RFG and the use of gasoline is a major contributor to ambient hydrocarbons measured in urban areas, the total NMHC should be expected to decrease. The regulations also call for specific reductions in benzene, thus the benzene concentration and weight percent in the atmosphere should be reduced. Additionally, the regulations call for reductions in the total aromatic hydrocarbon concentration. The PAMS measurements include several aromatic species, but may not adequately represent "total aromatics" when compared to fuel content. However, photochemical modeling studies by Ligocki and Yarwood (1994) have shown that when switching from conventional gasoline to RFG, the predicted reduction for C8 aromatic hydrocarbons (mostly xylenes) is 14 percent, while the predicted reduction in C9-C10 aromatic hydrocarbons is about 50 percent. A good surrogate for C8 aromatic hydrocarbons is total xylenes, which is measured by PAMS. A good surrogate for C9-C10 aromatic hydrocarbons is trimethylbenzene (TMB) which is also measured by PAMS. These modeling studies also predicted a decrease of 20 percent for 1,3-butadiene and 10 percent for formaldehyde. Both of these are measured at some of the PAMS sites.

Table ES-2.   Species (concentration and weight percent of NMHC) and ratios used as possible indicators of RFG implementation.

Indicator	Motivation
Benzene	Specific reductions called for by RFG regulations.
n-Butane	Normally reduced to lower RVP; RVP needs to be reduced for RFG.
1,3-Butadiene	Modeling studies predict a 20 percent decrease.
Xylenes	An aromatic compound; aromatics need to be reduced for RFG.
Formaldehyde	Modeling studies predict a 20 percent increase.
i-Butene	Decomposition product of MTBE; MTBE is an oxygenate used to meet RFG requirements.
Trimethylbenzenes (TMB)	An aromatic compound; aromatics need to be reduced for RFG, particularly C9-C10 aromatics.
C8 Alkanes	Possibly substituted for aromatics to meet RVP reductions.
NMHC	RFG should result in a reduction in the total amount of hydrocarbons.
Benzene/Toluene	If NMHC is not available, this is a good surrogate for the benzene weight percent since toluene is typically measured accurately and does not change as much with RFG as benzene.
n-Butane/i-Pentane	If NMHC is not available, this is a good surrogate for the n-butane weigh percent since i-pentane is not expected to change with RFG.
Benzene/Acetylene	If NMHC is not available, this is a good surrogate for the benzene weight percent since acetylene does not change with RFG and has few sources other than motor vehicle exhaust.
TMB/Toluene	If NMHC is not available, this is a good surrogate for the TMB weight percent since toluene is typically measured accurately and does not change as much with RFG as TMB.
TMB/Xylenes	TMB expected to change more than total xylenes.

The RVP reductions required by RFG may be met by the reduction of the butane content. Therefore, the trend in n-butane can be assessed. Unfortunately, the PAMS program does not require the measurement of the oxygen additives MTBE or ethanol. The olefin i-butene is a thermal decomposition product of MTBE, however, this hydrocarbon is not a PAMS target species and while it is reportedly measured at a few sites, it was rarely reported above the detection limit.

Previous investigations have shown that the species fractions (e.g., concentration of an individual hydrocarbon divided by the NMHC) show less variability than the concentrations (e.g., the interquartile ranges were typically smaller on a relative basis). Likely, the fluctuations caused by variations in meteorology (e.g., mixing depth) and emissions rates (e.g., traffic pattern changes) are greater on a concentration basis than on a weight fraction basis because these variations influence both the individual species and the NMHC.

It is important to consider that there are other sources of these hydrocarbons in ambient air besides motor vehicle emissions. For example, toluene is a commonly used solvent. However, most of the sites studied here were situated in areas with significant motor vehicle emissions. The emissions of these compounds usually build up during the morning rush hours while the mixing layer is shallow. These emissions react during the next several hours to produce the maximum downwind ozone concentrations during midday. The morning measurements at the PAMS sites (i.e., 0600-0900 local time) were used for the current analysis since mixing heights are low, concentrations are high, and photochemical reactivity is less important at this time of day.

ES.4.2 SPECIES RATIOS

The primary reason for investigating the changes in species ratios in this project is that not all data sets reported a total NMHC and thus the individual species data could not be normalized by computing a weight fraction. Instead, these data can be normalized by using the ratio of the concentrations of a species that is expected to change to the concentration of a species less changed and/or more stable in the atmosphere. Ratios of use in this study include:

- n-butane/i-pentane. Reductions in n-butane may be larger than reductions in i-pentane.

- Benzene/acetylene. Acetylene in the exhaust should not change significantly compared to benzene reductions.

- Benzene/toluene and benzene/xylenes. While toluene and xylenes are expected to decrease due to aromatic reductions, benzene reductions should be significantly larger.

- C9-C10 aromatic hydrocarbons/toluene and C9-C10 aromatic hydrocarbons/xylenes. Toluene and xylene reductions in motor vehicle exhaust may be modest compared to overall reductions in the aromatics.

Whenever possible, weight fractions were used in the analyses. However, the above ratios provided good surrogates when NMHC data were not available.

ES.5 RESULTS

In this project, the trends in more than 20 species concentrations, weight percents, or species ratios were investigated for every site. This investigation consisted of a review of annual box plots of the data and a statistical comparison of the annual data. Both nonparametric tests and t-tests were applied to the annual data. In most cases, the results were the same. For example, when the Kruskal-Wallis ANOVA indicated p-values less than 0.05, the t-test results also indicated p-values less than 0.05.

In general, the benzene weight percent and other benzene indicators (e.g., benzene/toluene ratio) showed statistically significant declines from 1994 to 1995. From 1995 to 1996, there were fewer statistically significant changes. Since much of the focus of RFG has been on the apparent ambient benzene reductions, **Table ES-3** summarizes the change in the ambient benzene weight percent from 1994 to 1995 and from 1995 to 1996. Significant decreases in summer morning ambient benzene weight percents were observed at most sites between 1994 and 1995 (data were not available at all sites). Some sites showed continued declines in ambient benzene levels in 1996 while other sites showed increases that offset some or all of the decreases from 1994 to 1995. For example, ambient benzene weight percent showed a statistically significant increase between 1995 and 1996 in Boston, which corresponds to OMS compliance survey data for the Boston area showing that fuel benzene content also increased.

Table ES-3.   Absolute changes in median ambient benzene weight percent between 1994/1995

and 1995/1996 (summer morning data). Only statistically significant changes are listed (i.e., p-values < 0.05 using the Kruskal-Wallis nonparametric test). Decreases are shaded. The 1994 median weight percent of benzene is provided as a point of reference.

Site	Median wt% Benzene 1994	Absolute Change 1994-1995	Absolute Change 1995-1996
Boston (Lynn), MA	2.53	-1.1	0.44
Washington (McMillan Reservoir), D.C.	3.49	-1.3	
E. Providence, RI	2.16	-0.38	-0.18
Milwaukee, WI	3.23	-0.40	0.40
Philadelphia, PA	3.06	-0.70	
Springfield (Chicopee), MA	1.99	-0.40	0.40
Los Angeles, CA	3.08	-0.80	-0.40[a]
Chicago, IL	2.89[b]	ns	-0.70
New York City (Bronx), NY	1.93[b]	ns	-0.70
Baton Rouge, LA[c]	2.02	-0.50	0.60
Harrington Beach, WI[c]	2.97	-0.44	1.1

[a] NIPER fuel data show an additional decrease in fuel benzene composition between 1995 and 1996.
[b] 1995 median value.
[c] RFG not implemented.
ns Data not suitable (e.g., cannot compute weight percent because of missing NMHC values).

There have been many documented changes in benzene levels in the fuel in California and declines in benzene have been noted in recent work (e.g., Hammond, 1996; Zielinska et al., 1997). In the data set available to this project, i-butene concentrations and weight fractions increased between 1994 and 1995; this is consistent with predictions for this hydrocarbon to increase with RFG (e.g., Ligocki and Yarwood, 1994). Decreases were observed in the weight fractions of n-butane, benzene, and xylenes, as well as the ratios of benzene/toluene, n-butane/i-pentane, and benzene/acetylene and the statistical tests confirmed the differences between the annual means and medians. All these changes are consistent with RFG introduction. Even though changes in California fuels were initiated in 1992, there was a change in the fuel from 1994 to 1995 and additional declines in fuel benzene in 1996 with the introduction of California Phase II RFG; the additional decline in ambient benzene weight percent, concentration, and related ratios observed in 1996 reflect these fuel changes.

At Baton Rouge, LA, a site in this study without RFG implementation, benzene weight percent and the ratios showed statistically significant differences (a decrease) between 1994 and 1995. However, there was a statistically significant increase in the benzene weight percent from 1995 to 1996 that was of a similar magnitude to the 1994 to 1995 decrease. Benzene concentrations showed little change. Also, while benzene weight percents showed a decline between 1994 and 1995, the benzene/acetylene and benzene/toluene ratios increased. This is not what was observed at the other sites discussed above where a consistent picture of benzene declines was typically exhibited (e.g., in most cases a decline in benzene weight percent or

concentration accompanied a decrease in the benzene ratios). Unfortunately, no fuel composition information was available for this city. Also, changes in non-motor vehicle local emissions may influence these results.

The analysis of the indicators selected to detect RFG suggests that RFG indeed influenced ambient concentrations. If it could be shown that a compound not expected to be reduced with RFG did not show a trend over the RFG implementation period, this would provide further evidence that changes in other compounds were due to the influence of RFG. Fortunately, isoprene, a hydrocarbon of primarily non-anthropogenic origin, is one of the PAMS target species and thus is available for this analysis. Isoprene emissions are a function of temperature and sunlight. Thus, isoprene concentrations and weight percent of total NMHC typically reach a maximum during midday or afternoon hours, and should vary day to day depending on the meteorological conditions. Peak isoprene concentrations have been found to vary widely from site to site depending upon the site's proximity to vegetation and other factors such as temperature, mixing height, and wind direction. For this analysis, the isoprene concentration and weight percent data at each site were investigated over the same time period as the RFG indicators (morning). There was no discernible trend in the isoprene concentrations over the period that RFG was implemented. Combined with the conclusion that the RFG indicators, such as the weight percent of benzene, did indeed decline over this same period, this analysis suggests that the trends in the RFG indicators were real and not due to other factors such as annual variations caused by meteorology.

ES.6 CONCLUSIONS

When one considers the above findings, a strong case emerges that supports the observation that a reduction of ambient benzene was related to the reduction of benzene in the fuel, and that there may also be significant reductions of other species such as aromatics. However, the results were not consistent across all of the sites. These inconsistencies may be due to inaccurate fuel data to document the expected changes with RFG implementation or with measurement problems at the PAMS sites.

1. INTRODUCTION

1.1 BACKGROUND

In accordance with the Clean Air Act Amendments of 1990, the use of reformulated gasoline (RFG) was required in the nine worst ozone nonattainment areas of the country. Other areas of the country also opted to implement RFG programs to supplement their clean air plans. The requirement to use RFG became effective at the retail level on January 1, 1995 outside California and in late 1992 in California.

In 1991, the National Academy of Sciences (NAS) released a report entitled, "Rethinking the Ozone Problem in Urban and Regional Air Pollution," which called for the establishment of monitoring networks to adequately track trends in ozone precursor emissions, corroborate emission inventories, and support photochemical modeling. The Photochemical Assessment Monitoring Stations (PAMS) program forms the basis of the EPA's response to recommendations from the National Academy of Sciences as well as the agency's response to the NAS report which is embodied in the Section 185b Report to Congress. Hydrocarbon and carbonyl data, as well as surface meteorology data, are collected at PAMS sites routinely. Many states implemented PAMS sites in 1994; similar data collection (PAMS-like) has been carried out in several areas for many years.

The Office of Mobile Sources (OMS) requested an evaluation of PAMS data to provide early confirmation of RFG's ambient air quality benefits. This report summarizes the objectives and tasks of this project, RFG use, the PAMS hydrocarbon database, and an assessment of trends observed in the ambient data.

1.2 OBJECTIVES AND TASKS

In this project, an evaluation and analysis of PAMS data was performed to provide some early confirmation of possible ambient air quality benefits attributable to the implementation of reformulated gasoline. Objectives of this project include the following:

- Determine if the PAMS measurements indicate real reductions in ambient concentrations of ozone precursors and toxics, rather than fluctuations.

- Determine to what extent these reductions in ambient concentrations can be linked to the implementation of the RFG program.

The tasks that were performed to meet these objectives included:

- Identify PAMS and other data sets to help address the project objectives.

- Determine that data were sufficient in quantity, quality, and scope to warrant further investigation at this time.

- Identify various ambient species and indicators to reflect changes in fuel composition due to the introduction of RFG. The ambient species and indicators include specific species concentrations, species mass fractions, and ratios of species concentrations.

- Determine sample size requirements for statistical determination of trends.

- Assess whether or not observed trends were statistically significant.

- Identify other auxiliary data sets that might also show changes due to the introduction of RFG (e.g., toxics data).

- Evaluate the consistency among the results using various species, indicators, auxiliary data sets, and various analysis methods. If the results are consistent, then a higher confidence level can be placed on the results; if some results are not consistent, then a lower confidence level can be placed on the results, and some investigative follow-up is needed.

- Recommend ways to expand and enhance the investigation, including recommendations on the appropriate analyses to perform when more data becomes available, on the availability and uses of additional auxiliary data sets, and on the use of more complex methods such as photochemical modeling in conjunction with data analysis.

1.3 ORGANIZATION OF THIS REPORT

The remainder of this report provides a discussion of reformulated gasoline implementation (Section 2), PAMS hydrocarbon data availability and validity (Section 3), ambient hydrocarbon trend analysis (Section 4), and a summary of conclusions and recommendations for further work (Section 5). Appendix A contains tables summarizing PAMS hydrocarbon data validation results. Appendix B contains tables summarizing the statistical test results from comparisons of annual data.

2. BACKGROUND AND TECHNICAL APPROACH

2.1 IMPLEMENTATION OF RFG

2.1.1 Motivation for RFG Implementation

To help meet clean air standards, the Clean Air Act Amendments of 1990 (CAAA) required the use of reformulated gasoline in the nine worst ozone nonattainment areas of the country. In addition, the CAAA allowed other areas of the country to "opt-in" to the reformulated gasoline program if those areas believed that RFG would be a cost-effective addition to their clean air plans. The requirement to use RFG became effective at the retail level on January 1, 1995. Today, approximately one quarter to one third of all gasoline fuel sold in the United States is RFG. RFG has achieved widespread use due to its significant potential as an ozone and toxic reduction measure. RFG use applies to fuel sold year round, and the fuel's performance standards require a minimum 15 percent reduction in hydrocarbon and toxics emissions. By the year 2000, a second phase of requirements will take effect and RFG must achieve at least a 20 to 25 percent reduction in hydrocarbons and toxics, and a 4 to 7 percent reduction in NO_x emissions.

2.1.2 RFG Content Requirements

This study focuses on the impact of the implementation of Phase I RFG on ambient concentrations. The principal differences between Phase I RFG and conventional gasoline include:

- Benzene levels are limited to no more than 1.0 percent by volume in RFG. In conventional gasoline, benzene can be as high as 5.0 percent by volume. Benzene is a proven human carcinogen and reductions in fuel benzene content constitute the bulk of the toxic reductions achieved through RFG implementation. Benzene is also an ozone precursor.

- A minimum 2.0 percent by weight of oxygen content is required in RFG. This requirement is typically met by either adding up to about 10 percent by volume ethanol or 11 percent by volume methyl tertiary butyl ether (MTBE). Oxygen is required by the CAAA to be included in RFG; ethanol and MTBE are the two most widely used oxygenates in RFG. Refiners are replacing aromatic hydrocarbons with high octane oxygenates to meet the oxygen content requirement and to lower aromatic hydrocarbon levels.

- The Reid Vapor Pressure (RVP) of the gasoline is reduced. Typically, butanes are reduced to decrease RVP. Higher alkanes and aromatics may be added to reduce fuel volatility. Reducing RVP reduces the fuel's volatility and constitutes the bulk of the hydrocarbon emission reductions achieved through RFG.

- Required reductions of air toxics and VOC emissions by 15 percent result in RFG fuels having lower aromatic content compared to baseline fuels. Typically, RFG fuels contain 20 to 27 percent aromatics whereas baseline fuel aromatics content was 32 percent (summer) and 26 percent (winter) by volume. Aromatic hydrocarbons include toxics such as benzene, toluene, and the xylenes; many of these compounds are also ozone precursors.

Table 2-1 highlights the major differences in fuel composition between conventional gasoline and RFG. Note that the information presented in Table 2-1 is a national summary provided by the EPA. Data available from industry shows similar differences between conventional fuel and RFG, although industry figures differ slightly in the estimate of average national fuel specifications.

Table 2-1. Conventional gasoline versus reformulated gasoline RFG (from EPA OMS web page http://www.epa.gov/OMSWWW/rfgnew.htm, updated by S. Romanow, OMS).

	Fuel Parameter Values (national basis)				
	Conventional Gasoline		Gasohol	Oxyfuel (2.7% wt)	VOC-Controlled ("Summer") Phase I RFG
	Average[a]	Range[b]	Average	Average	Average[e]
RVP[c] (PSI)	8.7-S 11.5-W	6.9-15.1	9.7-S 11.5-W	8.7-S 11.5-W	7.0-7.9[f]
T50 (°F)	207	141-251	202	205	199
T90 (°F)	332	286-369	316	318	325
Aromatics (volume %)	28.6	6.1-52.2	23.9	25.8	22.3
Olefins (volume %)	10.8	0.4-29.9	8.7	8.5	12.5
Benzene (volume %)	1.60	0.1-5.18	1.60	1.60	0.64
Sulfur (ppm)	338	10-1170	305	313	269
MTBE[d] (volume %)	--	0.1-13.8	--	15	11[g]
EtOH[d] (volume %)	--	0.1-10.4	10	7.7	10[h]

[a] As defined in the Clean Air Act.

[b] 1990 Motor Vehicle Manufacturer's Association (MVMA) survey.

[c] Winter (W) Reid Vapor Pressure (RVP) higher than Summer (S) to maintain vehicle performance.

[d] Oxygenate concentrations shown are for separate batches of fuel; combinations of both methyl tertiary butyl ether (MTBE) and ethanol (EtOH) in the same blend can never be above 15 volume percent total.

[e] Except as noted, values in this column are volume-weighted average based on preliminary analysis of a portion of 1996 RFG batch reporting data submitted to EPA by refineries and importers. Numbers are subject to change.

[f] The lower RVP number is average for RFG designated for sale in VOC Control Region 1; the higher RVP number for VOC Control Region 2. In general, Region 1 areas are to the south and/or west.

[g] This is a "typical" value for an RFG batch containing MTBE as the sole oxygenate; this provides about 2% oxygen by weight.

[h] This is a "typical" value for an RFG batch containing ethanol. Ethanol is not combined with other oxygenates in VOC-controlled RFG. There are economic incentives for blending ethanol at this level, which provides about 3.5% oxygen by weight.

2.1.3 Implementation Milestones

The federal RFG requirement has two key phase-in milestones: Phase I RFG was required to be available at gasoline retail operations beginning January 1, 1995. Phase II RFG, which will require further hydrocarbon and toxic reductions, is required to be available in the year 2000. In addition, California has had separate fuel requirements that also require gasoline reformulation, although California's fuel requirements differ somewhat from federal RFG mandates. Key milestones for both California and federal RFG include:

- November 1, 1992: implementation of California's Phase 1 gasoline.
- January 1, 1995: federal RFG (EPA Phase I) required for sale at retail level in nation's nine worst ozone nonattainment areas.
- March 1, 1996: California's Phase 2 RFG required at refinery level.
- April 15, 1996: California's Phase 2 RFG required at terminal level.
- June 1, 1996: California's Phase 2 RFG required at retail level (entire state).
- January 1, 1998: federal RFG moves from the Simple Model, which tracks four fuel parameters (RVP, benzene, oxygenates, and aromatics), to the Complex Model, which tracks three additional parameters (sulfur, olefins, and distillation range limitations).
- January 1, 2000: federal RFG (EPA Phase II) goes into effect in areas previously covered by Phase I.

2.1.4 RFG Implementation Areas

The CAAA required nine areas in the nation to sell only RFG year-round after January 1, 1995. These areas included:

1. Baltimore, MD, and six surrounding counties.
2. Chicago/Gary, IN, including eight full or partial Illinois counties and one county in Indiana.
3. Hartford, CT, including the cities of New Britain, Middletown, New Haven, Meriden, and Waterbury, plus parts of six surrounding counties.
4. Houston/Galveston, TX, including eight counties.
5. Los Angeles, Anaheim, and Riverside, CA, including Los Angeles, Ventura, and Orange counties and parts of San Bernardino and Riverside counties.
6. Milwaukee/Racine, WI, including six counties.
7. New York metropolitan area, including New York City, Long Island, four upstate New York counties, 12 northern NJ counties, and Fairfield County, CT.
8. Philadelphia, PA, including two counties in Delaware, one in Maryland, six in New Jersey, and five in Pennsylvania.
9. San Diego County, CA.

Other areas voluntarily "opted in" to the RFG program. These areas included:

1. Connecticut - those parts of the state not included in the Hartford or New York City areas.
2. Delaware - Sussex County.
3. District of Columbia.
4. Kentucky - Boone, Campbell, and Jefferson counties and parts of Bullitt and Oldham counties.
5. Maine - Androscoggin, Cumberland, Lincoln, Kennebec, Knox, Sagadahoc, and York counties.
6. Maryland - Calvert, Charles, Frederick, Montgomery, Prince, Georges, Queen Anne's, and Kent counties.
7. Massachusetts - entire state.
8. New Hampshire - Hillsborough, Merrimack, Rockingham, and Strafford counties.
9. New Jersey - Atlantic, Cape May, and Warren counties.
10. New York - Duchess County and parts of Essex County.
11. Rhode Island - entire state.
12. Texas - Collin, Dallas, Denton, and Tarrant counties.
13. Virginia - 17 independent cities, primarily in the Washington, Richmond, and Tidewater areas.

In June 1996, the EPA finalized a rule-making that allowed several states to "opt out" of the RFG program. These areas included portions of the following states (RFG had not yet been implemented in the areas that "opted out"): Maine (Hancock and Waldo counties), Pennsylvania (Allentown, Altoona, Erie, Harrisburg, Johnstown, Lancaster, Pittsburgh, Scranton, York, and Reading metropolitan areas), Ohio (Youngstown metropolitan areas), New York (Albany and Buffalo metropolitan areas and Jefferson County), and Wisconsin Sheboygan, Manitowoc, and Kewaunee areas).

2.1.5 Other Fuel Requirements and Actual RFG Implementation

2.1.5.1 Motor Vehicle Emissions and Fuel Effects

A major focus of the CAAA was reformulating fuel to achieve more complete combustion. To reduce carbon monoxide and ozone, and to cut other air toxics, the CAAA required the addition of oxygen to gasoline. Oxygenates added included either alcohols or ethers, including MTBE, ethanol, and ethyl tertiary butyl ether (ETBE). RFG must achieve 15 percent reductions in smog producing emissions during the summer and must produce no net increase in nitrogen oxides. The use of oxygenates in gasoline allows petroleum refiners to decrease the use of toxic substances such as benzene, 1,3-butadiene, toluene, and xylenes, and still produce a high quality engine fuel. Specifically, the fuel parameters regulated under Federal Phase I include RVP, oxygen content, and benzene. However, the EPA cannot prescribe how each refiner will achieve the RFG goals. As a result, the composition of RFG varies throughout the country, depending on the sales volume provided by each refiner in each market.

It is important to remember that year-to-year changes in ambient air quality can be related to many variables including real changes in emissions (from both mobile and non-mobile sources) and differences in meteorology. In addition, emission changes are also expected from other control programs on stationary and mobile sources (including Inspection and Maintenance - I/M and Stage II vapor recovery at gasoline refueling stations) as well as emission reductions expected from fleet turnover. Several published reports describe the expected emission changes due to RFG on both exhaust and evaporative emissions. But it is important to note that some of the emission effects are directly tied to the type and amount of oxygenate in the fuel and can only be accurately computed from the actual fuel composition. Further it is also important to note that while the goals of the RFG program include reductions in key emissions of ozone forming chemicals, RFG is expected to increase emissions of some compounds. Increases in emissions relative to pre-RFG fuels are expected for compounds added to the fuel (e.g., MTBE, ETBE, EtOH) and for by-products, such as formaldehyde, from combustion of the fuel additives.

2.1.5.2 Fuel Survey Data

Independent of fuel changes related to RFG, some areas of the country have modified gasoline's RVP and oxygenate content. Oxygenate use is required in some areas of the country to help reduce wintertime carbon monoxide concentrations, and RVP regulations have been in place to help reduce hydrocarbon emissions that contribute to ozone formation. Therefore, a review of RFG's effectiveness needs to assess what best represents the "before and after" conditions for gasoline in the region to be studied (i.e., what was the gasoline's composition before and after RFG implementation). In addition, fuel characteristics vary from one region of the country to another, so fuel characteristics need to be specific to each area of study.

A significant effort was made to track down fuel content before and after RFG implementation on a city-by-city basis. Three data sets proved useful for this investigation:

- American Automobile Manufacturers Association (AAMA, previously Motor Vehicle Manufacturers Association, MVMA) fuel survey results for 1994 and 1995.

- National Institute for Petroleum and Energy Research (NIPER) summer gasoline analyses for 1993-1996.

- RFG Survey Association fuel survey data for 1995 and 1996 collected at the retail level.

The AAMA fuel survey results for 1994 and 1995 are summarized in **Tables 2-2 and 2-3** for six cities. Most of the metropolitan areas investigated in this study showed reductions in RVP, aromatic content, olefin content, and benzene with increases in the saturate (paraffin) content.

Table 2-2. Average fuel composition (percent by volume) as reported for 1994 and 1995 in several metropolitan areas (source = AAMA fuel survey results, provided by S. Romanow, OMS). Regular unleaded fuel; RFG in 1995.

City	Year	RVP[a]	Aromatic	Olefins	Saturates	Benzene	MTBE
Boston, MA	94	8.5	27	18	55	1.5	1.6
Chicago, IL	94	9.1	28	11	61	1.4	0.1[b]
Los Angeles, CA	94	7.5	32	9	59	1.4	0.5
New York, NY	94	8.7	24	23	52	1.1	0.8
Philadelphia, PA	94	8.6	29	18	54	0.9	1.4
Washington, D.C.	94	7.5	29	15	56	0.8	1.0
Boston, MA	95	7.8	25	8	67	0.5	10.2
Chicago, IL	95	8.0	27	10	63	0.9	<2.6[c]
Los Angeles, CA	95	7.3	29	9	62	0.8	11.1
New York, NY	95	8.1	26	16	58	0.7	10.7
Philadelphia, PA	95	8.1	28	15	62	0.7	11.5
Washington, D.C.	95	7.2	30	15	55	0.7	11

[a] PSI
[b] Ethanol content less than 4.8 percent. All other cities less than 0.1 percent ethanol.
[c] Ethanol content less than 6.4 percent. All other cities less than 0.1 percent ethanol.

Table 2-3. Absolute change in average fuel composition between 1995 and 1994 (AAMA fuel survey results). Regular unleaded fuel.

City	RVP[a]	Aromatic	Olefins	Saturates	Benzene	MTBE
Boston, MA	-0.7	-2.2	-10	12	-1.0	8.6
Chicago, IL	-1.1	-0.9	-0.9	1.8	-0.5	
Los Angeles, CA	-0.2	-3.1	-0.3	3.4	-0.6	10.6
New York, NY	-0.6	2.0	-7.9	6.0	-0.4	9.9
Philadelphia, PA	-0.5	-1.0	-2.8	3.8	-0.2	10.1
Washington, D.C.	-0.3	1.2	-0.4	-0.8	-0.1	10.0

[a] PSI

Depending on the availability of ambient data, the larger changes in fuel composition will probably be easier to discern (e.g., easier in Boston compared to New York).

The NIPER survey data for 1993-1996 were also investigated (**Tables 2-4 and 2-5**). These surveys are provided on a regional basis in the report although data from individual cities are available at a modest cost (but not obtained). From a qualitative comparison, the NIPER data were consistent with the AAMA data.

Table 2-4. Fuel survey results published by NIPER for 1993-1996; unleaded fuel with 88 to 89.9 knock index. Benzene and MTBE values given in percent by volume. Vapor pressure (VP) at 100°F, psi.

Unleaded Gasoline

	1993			1994			1995			1996		
	Benzene	VP	MTBE	Benzene	VP	MTBE	Benzene	VP	MTBE	Benzene	VP	MTBE
Northeast	1.14	8.3	1.4	1.28	8.5	2.8	0.88	8.7	1.9	1.18	8.5	1.2
Mid Atlantic	0.84	7	2.1	0.97	7.3	2.7						
Southeast	1.05	7.3	1.1	1.22	7.6	1.5	1.01	7.5	1	.85	7.8	.7
North central	1.22	8.3	2.1	1.57	8.5	0.5	1.36	8.6	1.9	1.2	8.1	.8
Eastern Texas	1.19	7.7	0	1.82	7.9	1.3	1.3	8.2	1.3	1.3	8.3	1.2
Southern Cal	1.41	7.3	3.2	1.22	7.6	1.5						

Gasohol

	1993			1994			1995 RFG			1996 RFG		
	Benzene	VP	Ethanol	Benzene	VP	Ethanol	Benzene	VP	Ethanol	Benzene	VP	Ethanol
Northeast				1.36	9.4	9.5						
Southeast	0.65	9	10.2	1.1	8.6	9.4						
North central	1.18	9.1	10.8	1.4	9.6	10.5	0.72	7.9	7.7	.82	7.8	10.4
Eastern Texas	0.8	8.7	10	3.01	9.7	9.6						

Reformulated Gasoline

	1995			1996		
	Benzene	VP	MTBE	Benzene	VP	MTBE
Northeast	0.63	7.9	9.9	.67	7.9	11.1
Mid Atlantic	0.61	7.1	10.3	.65	7.1	9.3
North Central	0.82	8	11.1			
Eastern Texas	0.75	7	9.8	.67	7.1	9.7
Southern Cal	0.79	7.2	11.4			

Northeast = ME, MA, NJ, NY, PA, RI
Mid-Atlantic Coast = D.C., MD, VA
North Central = IL, IN, MI, MN
Southeast = AL, AR, GA, LA, NC, SC
Southern California = Los Angeles, San Diego
Eastern Texas = OK, Houston, Dallas
Not all sites in the northeast, north central, or Texas implemented RFG.

Table 2-5. Differences between 1995 and 1994 (NIPER 1993-1996; unleaded fuel with 88 to 89.9 knock index). Benzene and MTBE values given in percent by volume. Vapor pressure (VP) at 100°F, psi.

	Gasoline (1995 RFG-1994)			Gasohol (1995-1994)		
	Benzene	VP	MTBE	Benzene	VP	Ethanol
Northeast	-0.65	-0.6	7.1			
Mid Atlantic	-0.36	-0.2	7.6			
North Central	-0.75	-0.5	10.6	-0.68	-1.7	-2.8
Eastern Texas	-1.07	-0.9	8.5			
Southern Cal	-0.43	-0.4	9.9			

	RFG Gasoline (1996 -1995)			RFG Gasohol (1996-1995)		
	Benzene	VP	MTBE	Benzene	VP	Ethanol
Northeast	0.04	0	1.2			
Mid Atlantic	0.04	0	-1.0			
North Central				0.1	-0.1	2.7
Eastern Texas	0.08	0.1	-0.1			
Southern Cal	-0.31	-0.2	-0.2			

Northeast = ME, MA, NJ, NY, PA, RI
Mid-Atlantic Coast = D.C., MD, VA
North Central = IL, IN, MI, MN
Southeast = AL, AR, GA, LA, NC, SC
Southern California = Los Angeles, San Diego
Eastern Texas = OK, Houston, Dallas
Not all sites in the northeast, north central, or Texas implemented RFG.

These data were useful for ascertaining fuel changes from 1994 to 1995 in most cities. However, there were a few shortcomings, as follows:

- Data are not available for each city with a PAMS type II monitor (defined in Section 3.1) that were investigated in this report, including E. Providence, RI, Baton Rouge, LA, and Milwaukee, WI.

- The NIPER data would be more representative if it were averaged by sales volume. Only average data for the samples taken in each region are provided.

- More information is needed regarding implementation dates in some of the areas.

The RFG compliance survey data compiled by the RFG Survey Association are provided for 1995 and 1996 in **Tables 2-6 and 2-7**. Table 2-6 shows levels of benzene, aromatics, RVP, and total oxygen for the VOC-control season only for nine of the study areas for 1995 and 1996. The percent change between 1995 and 1996 are provided in Table 2-7; only statistically significant changes at the 95 percent confidence level are provided. Benzene levels in fuel showed some increases from 1995 to 1996 in the Boston, Washington, DC, Milwaukee, and Chicago areas. Aromatics levels in the fuel showed increases in Boston, Rhode Island, Philadelphia, and New York areas and decreases in Springfield and Washington, DC.

2.2 TECHNICAL APPROACH

2.2.1 Useful Indicator Species and Ratios

As a part of the Coordinating Research Council (CRC) Model Evaluation Feasibility Study (Stoeckenius et al., 1995), STI investigated toxic species, hydrocarbon, and carbonyl compound data collected in the South Coast Air Basin (SoCAB) of California from 1990 through 1993 (Main and Roberts, 1994). Using these data, the temporal trends of selected species, species groups, and ratios were evaluated to assess the usefulness of these parameters as indicators of change in motor vehicle fuel composition (Main and Roberts, 1994; Main et al., 1995). Based on predicted changes in the evaporative and exhaust emissions due to the introduction of California Phase I RFG (Ligocki and Yarwood, 1994) and our knowledge of the ambient data quality and availability, this investigation focused on several species, species groups, and ratios (if available). Consensus among several indicators gives us more confidence in the study conclusions. The useful indicators are discussed below.

2.2.1.1 Concentrations and Weight Fractions

There are several species and species group concentrations that may be good ambient indicators of fuel RFG changes. For example, since 15 percent reductions in smog-reducing emissions were required with RFG, the total nonmethane hydrocarbon (NMHC) should be expected to decrease. The regulations also call for specific reductions in benzene, thus the

Table 2-6. 1995 and 1996 RFG Compliance Survey Data (provided by S. Romanow, OMS).

Area	Year	Benzene (vol%)	Aromatics (vol%)	RVP (psi)	Total Oxygen (Wt%)
Boston-Worcester, MA	1995	0.57	22.2	7.8	2.18
Washington, DC	1995	0.60	27.2	7.0	2.14
Rhode Island	1995	0.62	19.8	7.8	2.26
Milwaukee-Racine, WI	1995	0.93	25.0	7.9	2.61
Phila.-Wilm, DE-Trenton, NJ	1995	0.64	23.7	7.9	2.09
Springfield, MA	1995	0.60	22.4	7.8	2.17
Chicago-Lake Co., IL, Gary, IN	1995	0.83	24.8	7.9	2.59
NY-NJ-Long Is.-CT	1995	0.58	23.2	7.9	2.08
Hartford, CT	1995	0.63	23.3	7.9	2.12
Boston-Worcester, MA	1996	0.64	23.8	7.8	2.12
Washington, DC	1996	0.63	26.1	7.1	1.92
Rhode Island	1996	0.57	23.3	7.8	2.06
Milwaukee-Racine, WI	1996	1.0	24.4	7.9	3.42
Phila.-Wilm, DE-Trenton, NJ	1996	0.64	25.5	7.9	2.00
Springfield, MA	1996	0.59	20.7	7.8	2.12
Chicago-Lake Co., IL, Gary, IN	1996	0.88	24.1	7.9	3.36
NY-NJ-Long Is.-CT	1996	0.61	24.3	7.9	1.99
Hartford, CT	1996	0.64	23.7	7.8	2.02

Table 2-7. Percent change in fuel properties from 1995 to 1996 as reported in the RFG compliance survey data.

Area	Benzene	Aromatics	RVP	Total Oxygen
Boston-Worcester, MA	13.4	7.3		-2.8
Washington, DC	5.9	-4.3		-10.3
Rhode Island		18.0		-8.9
Milwaukee-Racine, WI	10.3			31.0
Phila.-Wilm, DE-Trenton, NJ		7.8		-3.9
Springfield, MA		-7.5		-2.1
Chicago-Lake Co., IL, Gary, IN	6.6			29.8
NY-NJ-Long Is.-CT		4.9		-4.3
Hartford, CT			-0.4	-4.8

Negative number indicates decrease in 1996.
Only statistically significant changes provided.

benzene concentration and weight percent in the atmosphere should be reduced. Additionally, the regulations call for reductions in the total aromatic hydrocarbon concentration. The PAMS measurements include several aromatic species, but may not adequately represent "total aromatics" when compared to fuel content. However, photochemical modeling studies by Ligocki and Yarwood have shown that predicted reduction when switching from conventional gasoline to RFG for C8 aromatic hydrocarbons (mostly xylenes) is 14 percent, while the predicted reduction in C9-C10 aromatic hydrocarbons is about 50 percent. A good surrogate for C8 aromatic hydrocarbons is total xylenes, which is measured by PAMS. A good surrogate for C9-C10 aromatic hydrocarbons is trimethylbenzene (TMB) which is also measured by PAMS. These modeling studies also predicted a decrease of 20 percent for 1,3-butadiene, and 10 percent for formaldehyde. Both of these are measured at a few of the PAMS sites. Discussions with analysts in California indicate that concentrations and weight fractions of 2-methylheptane (a C8 alkane) added to the California fuels in the past couple of years (Poore, 1997) may increase because refiners use it as an aromatic hydrocarbon substitute.

The RVP reductions required by RFG may be met by the reduction of the butane content. Therefore, the trend in n-butane will be assessed. Unfortunately, the PAMS program does not require the measurement of the oxygen additives MTBE or ethanol. The olefin i-butene is a thermal decomposition product of MTBE, however, this hydrocarbon is not a PAMS target species and while it is reportedly measured at a few sites, it was rarely reported above the detection limit.

Previous investigations have shown that the species fractions (e.g., concentration of an individual hydrocarbon divided by the NMHC) show less variability than the concentrations (e.g., the interquartile ranges were typically smaller on a relative basis). Likely, the fluctuations caused by variations in meteorology (e.g., mixing depth) and emissions rates (e.g., traffic pattern changes) are greater on a concentration basis than on a weight fraction basis because they influence both the individual species and the NMHC. For these reasons, when NMHC data are available, weight fractions will be calculated and used for the trend analysis.

It is important to consider that there are other sources of these hydrocarbons in ambient air besides motor vehicle emissions. For example, toluene is a commonly used solvent. However, most of the sites studied here were situated in areas with significant motor vehicle emissions.

The emissions of these compounds usually build up during the morning rush hours while the mixing layer is beginning to form. These emissions react during the next several hours to produce the maximum downwind ozone concentrations during midday. The morning concentrations measured at the PAMS sites will be used for the trend analysis since mixing heights are low, concentrations are high, and photochemical reactivity is less important at this time of day.

2.2.1.2 Species Ratios

The primary reason for investigating the changes in species ratios in this project is that not all data sets reported a total NMHC and thus the individual species data could not be normalized by computing a weight fraction. Instead, these data can be normalized by using the ratio of the concentrations of a species that is expected to change to the concentration of a species less changed and/or more stable in the atmosphere. Ratios of use in this study include:

- n-butane/i-pentane. Reductions in n-butane may be larger than reductions in i-pentane.

- Benzene/acetylene. Acetylene in the exhaust should not change significantly compared to benzene reductions.

- Benzene/toluene and benzene/xylenes. While toluene and xylenes are expected to decrease due to aromatic reductions, benzene reductions should be significantly larger.

- C9-C10 aromatic hydrocarbons/toluene and C9-C10 aromatic hydrocarbons/xylenes. Toluene and xylene reductions in motor vehicle exhaust may be modest compared to overall reductions in the aromatics.

Whenever possible, weight fractions were used in the analyses. However, the above ratios provided good surrogates when NMHC data were not available.

2.2.2 Formulation of the Statistical Test

After selection of indicators, the next step is to test whether there were statistically significant changes in the indicators selected between the periods before and after RFG was implemented. It is important to test whether the mean values of these ratios decline from the period before to the period after RFG implementation. This test might include the average for 1994 compared to 1995, or the average of 1993 and 1994 compared to the average of 1995 and 1996. To illustrate the statistical concepts used for detecting the trend, the ambient weight percent of benzene will be used. Also, the morning concentrations were used because the morning rush hour is the period when ozone precursors accumulate in the atmosphere leading to the highest precursor concentrations and it is more likely that changes will be observed when concentrations are high than when concentrations are low.

When formulating a statistical test, the standard procedure is to develop a null (H_o) and alternative hypothesis (H_1). Letting m_x equal the true mean of the ambient benzene weight percent before RFG and m_y equal the true mean weight percent of ambient benzene after the implementation of RFG, the null and alternative hypotheses for this comparison are as follows:

$$H_o: m_x = m_y$$
$$H_1: m_x \neq m_y$$

The true means (m_x and m_y) represent the actual mean concentration over the sampling period free from any biases. From the data, only the sample means ($\overline{X}$ and $\overline{Y}$) can be obtained. Since there is error associated with the measurement of the benzene concentration, the sample means may be different than the true means. One of the principal reasons that a large sample size is needed is to reduce any potential bias created by the measurement error. The variability in the annual mean concentrations associated with meteorology will be addressed in later analyses. However, the day-to-day variability in the benzene weight percent associated with meteorology, which also has a large effect, is built into this calculation, since it is indistinguishable from the measurement bias.

The next step is to develop an appropriate statistical test for the above hypothesis. The *t*-statistic for the comparison of two means is as follows:

$$t = \frac{\overline{X} - \overline{Y}}{S_p \sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$$

where *n* and *m* are the sample sizes and S_p is the pooled standard deviation. If this *t*-statistic is statistically significant (e.g., with the probability of means being different above 95 percent), the null hypothesis will be rejected and the alternative hypothesis accepted. Thus, for this example, once can conclude that there is a statistically significant trend for the weight percent of benzene before and after RFG. Conversely, if the *t*-statistic is not significant (e.g., with the probability of the means being different below 95 percent), the null hypothesis will be accepted and the alternative hypothesis rejected. Thus, one can conclude that there was no reduction (at least below the minimum detectable reduction determined from the sample size calculation in the next section) in weight percent benzene before and after RFG.

As an example, the t-statistics were calculated for the ambient benzene weight percent at 0600 EST for 1994 and 1995 at McMillan Reservoir, D.C. and Lynn, Massachusetts. **Table 2-8** presents the mean, standard deviation, sample size, and t-statistic for both sites. The means in 1994 were higher than the means in 1995 in all three cases, suggesting that the benzene weight percent was reduced as the result of the implementation of Phase I RFG in 1995. Generally, for these sample size values, the t-statistic must be above 2 for the means in 1994 and 1995 to be statistically different (i.e., rejecting the null hypothesis) at the 95th percentile confidence range. The t-values for Lynn, MA and McMillan Reservoir, D.C. easily meet this criterion.

An alternative, non-parametric test is the Mann-Whitney U Test. This test is analogous to the t-test except that it is based on the ranks of the values within each year, not the actual values. Thus, it is essentially a test based on the median values that does not assume any underlying distribution for the data.

Table 2-8. Results of the comparison of mean ambient benzene weight percents at 0600 EST during 1994 and 1995 at McMillan Reservoir, D.C. and Lynn, MA. The t-values > 2 indicate statistically significant differences between 1994 and 1995 data sets.

City	Year	N	Mean	Standard Deviation	t-value
Lynn, MA	1994	32	2.7	0.8	5.1
	1995	92	1.7	1.0	
McMillan Reservoir, D.C.	1994	30	3.5	0.6	8.1
	1995	81	2.2	0.8	

Note: The tests for Lynn, MA and McMillan Reservoir, D.C. were done for 1-hr samples from 0600-0700 EST.

2.2.2.1 Sample Size Considerations

The determination of the necessary sample size to detect the hypothesized trend is referred to as a power calculation. Since, in this case, the sample size is already fixed (historical data is being used), the results of the power calculation will give the minimum trend that can be statistically detected given the sample size. If the true trend is below this value, it will not be statistically undetectable. In other words, a null result for the statistical test (i.e., no detectable trend) will mean that there is either no true trend or that the true trend is too small to be statistically detectable given the sample size and variance. To properly interpret our results (particularly if there is a null result), it is necessary to have an approximation of the minimum detectable trend.

When comparing two means, there are two potential errors one can make:

(1) Type I error - rejecting the null hypothesis when it is in fact true, or
(2) Type II error - accepting the null hypothesis when it is in fact false.

The possibility of making both of these errors needs to be accounted for in the sample size calculation. Typically, researchers designing a study stipulate that the probability of making a Type I error (commonly known as a) be less than 5 percent, and the probability of making a Type II error (commonly known as b) be less than 20 percent. These standard assumptions will be used in this example.

The derivation of the formulas used to determine an adequate sample size for this problem are complicated and are discussed elsewhere (see Walpole and Myers, 1989). However, standard statistical tables are available to produce these calculations. Typically, these tables include four variables: (1) a (allowable Type I error), (2) b (allowable Type II

error), (3) the sample size, and (4) a variable, Δ. Accordingly, Δ is defined as:

$$\Delta = \frac{d}{s}$$

where, d, is the difference between the two means, as follows:

$$d = m_x - m_y$$

and s is the standard deviation of m_x and m_y, which are assumed to be equal. In other words, Δ is the ratio of the difference in means and the standard deviation. The exact solution to the problem with different variances is very complicated (Walpole and Myers, 1989) and well beyond the needs of the calculation to provide a reasonable estimate of the minimum detectable difference. However, for the values in Table 2-8, a test of the equality of variances (the F-test) showed that they were statistically indistinguishable.

In this study, solving for d will give the minimum difference in the mean concentrations that can be detected. Returning to the benzene weight percent examples, one can calculate the minimum difference in the benzene weight percent that can be statistically differentiated. For both Lynn and McMillan Reservoir, there were considerably more samples available for 1995 compared to 1994. As mentioned above, the calculation of sample size for a comparison of means with different sample sizes is very complicated. Therefore, for simplicity, the calculation will be done for both the sample sizes in 1994 and 1995. These results are summarized in **Table 2-9**. The t-tests done for the examples above showed that the means were statistically different. Therefore, it was expected that the minimum detectable differences would be lower than the actual differences. The calculation shows that with the higher sample sizes (i.e., this assumes that there were as many samples in 1994 as there were in 1995), the minimum detectable difference was between 0.2 and 0.3. The minimum detectable differences for the lower sample sizes were 0.3-0.5.

As a rough approximation, it is reasonable to assume that the weight percent changes in the atmosphere are commensurate with the weight percent changes in the fuel composition. In this case, according to Table 2-3, either sample size would be sufficient to detect the change in the benzene weight percent in Boston, Chicago, and Los Angeles. For New York and Philadelphia, the higher sample size might be needed. For Washington, DC, the change might be too small to detect. Furthermore, many of the compounds listed in Section 2.2.1 as potential indicators of RFG have lower concentrations than benzene. Therefore, we should expect that some of the null results for the statistical tests of these compounds simply means that there is not enough statistical power to detect the differences. However, it should be noted that this estimate is based on rough assumptions and is only designed to provide a general approximation of the influence of sample size on the calculation.

Table 2-9. Summary of the minimum detectable difference in benzene weight percent at Lynn, MA and McMillan Reservoir, D.C. The calculation assumes a Type I error of 5 percent, and a Type II error of 20 percent and uses the pooled standard deviation.

Site	Sample Size	d	Minimum Detectable Difference
Lynn, MA	32	0.46	0.44
	92	0.27	0.26
McMillan Reservoir, D.C.	30	0.47	0.35
	81	0.26	0.20

2.3 PRELIMINARY RESULTS FROM OTHER STUDIES

The EPA Office of Air Quality Planning and Standards (OAQPS) has performed preliminary analyses of PAMS data averaged across all sites for 1994-1995 to investigate potential trends in important VOCs (U.S. EPA, 1996). The most noticeable change in the two-year period studied was a decline in benzene concentration. Since RFG has reduced benzene content, the decline may be attributable, at least in part, to RFG implementation. In fact, OAQPS found that more sites in RFG areas showed statistically significant benzene reductions than areas without RFG. Ambient benzene reductions have also been noted in California (Hammond, 1996).

Zielinska et al. (1997) reported on ambient hydrocarbon concentrations in Los Angeles during the summers of 1995 (introduction of Federal Phase I RFG) and 1996 (introduction of California Phase II RFG). In this study, ambient concentrations of hydrocarbons as well as MTBE and carbonyls were investigated. Significant reductions were observed in ambient concentrations and weight fractions of C6 olefins, C9 and C10 aromatic hydrocarbons, and benzene. Significant increases were observed in i-butene and MTBE. Formaldehyde showed no change.

3. PAMS HYDROCARBON DATA AVAILABILITY AND QUALITY CONTROL

3.1 PHOTOCHEMICAL ASSESSMENT MONITORING STATIONS (PAMS)

In 1991, the National Academy of Sciences (NAS) released a report entitled, "Rethinking the Ozone Problem in Urban and Regional Air Pollution," which criticized the EPA for failing to establish monitoring networks to adequately track trends in ozone precursor emissions, corroborate emission inventories, and support photochemical modeling. The PAMS program forms the basis of the EPA's response to recommendations from the National Academy of Sciences. The PAMS program required the establishment of an enhanced monitoring network in all ozone nonattainment areas classified as serious, severe, or extreme (U.S. EPA, 1996). These areas include: *extreme* - the South Coast Air Basin (SoCAB) of California (the Los Angeles area); *severe* - Southeast Desert Modified air quality management area, Baltimore, MD; Chicago, IL/Gary, IN/Milwaukee, WI; Houston/Galveston, TX; New York/New Jersey/Long Island (NY-NJ-CT); Philadelphia/Wilmington/Trenton (PA-NJ-DE-MD); San Diego, CA; Ventura County, CA; and *serious* - Atlanta, GA; Baton Rouge, LA; Beaumont-Port Arthur, TX; Boston-Lawrence-Worchester, MA-NH; Greater Connecticut; El Paso, TX; Portsmouth-Dover-Rochester, NH-ME; Providence-Pawtucket-Fall River, RI-MA; Sacramento, CA; San Joaquin Valley, CA; Springfield, MA; and Washington, D.C.-MD-VA. Most of these metropolitan areas also have implemented RFG, however, not all of these areas have chosen to implement the RFG program, including Atlanta, Baton Rouge, and El Paso.

The PAMS network consists of as many as five monitoring stations depending on the area's population (U.S. EPA, 1994a). Site locations are based upon careful consideration of the meteorology, topography, and proximity to emissions of ozone precursors (VOC and nitrogen oxides). There are up to four types of PAMS sites based on the site location relative to the emissions and transport pathways in an area, defined as follows:

- Type I - located to provide upwind and background characterization of ozone and precursors being transported into an area.

- Type II - located to document the maximum ozone precursor emissions impact site. These sites are typically located downwind of the central business district and operate on the most intensive monitoring schedule of the PAMS sites. The proximity of the sites to the emissions sources in an area, and the frequency of the data collection, make these sites probably the best suited for investigating the effects of RFG on ambient VOC concentrations and composition.

- Type III - located to measure the maximum ozone concentration and typically sited further downwind than the type II sites.

- Type IV - located downwind of the nonattainment areas to assess the extreme downwind conditions. In some parts of the country, such as the Northeast urban corridor, a PAMS site may be defined as both a type I and type IV site depending upon its location.

Hydrocarbon and carbonyl data, as well as the surface meteorological data (e.g., temperature, wind speed, and wind direction) collected at the PAMS sites are available through the EPA's Aerometric Information Retrieval System (AIRS). According to the National Trends Report (U.S. EPA, 1996), there were 37 PAMS sites operating in 1994, 54 in 1995, and 65 in 1996. The National Trends report noted that most of the PAMS sites reported their data to AIRS; however, in past analyses, the 1994 data were not as readily available. Likewise, data from the summer of 1996 are not available from all sites. Note that most PAMS sites are operated only during the ozone season, which typically includes May or June through August or September.

Many of the PAMS sites employ an automated gas chromatography system (auto-GC) which reports hydrocarbon concentrations of 57 target hydrocarbon species and the total nonmethane hydrocarbon (NMHC) on an hourly basis. At other sites, canisters are collected for analysis. At most of these sites, eight 3-hr samples are collected either every day or every third day depending upon the site. Analyses of the canisters are performed using the TO-12 method; results for the PAMS target compounds are provided. Carbonyl compounds were collected at the PAMS sites using commercially available DNPH-impregnated silica gel cartridges. Results from most of the PAMS sites are reported only for formaldehyde, acetaldehyde, and acetone. Not all sites have reported their carbonyl results.

Two other issues are important when considering the PAMS data. First, at some PAMS sites, collection and/or analysis methods may have changed during the time period 1994-1996; this will need to be considered in any analysis of data at that site. For example, 3-hr canister data were collected at the Bronx, NY site during the summer of 1994, while during 1995 and 1996, the state operated an auto-GC at the site and reported hydrocarbon concentrations every hour. Second, the data completeness may vary at the sites from year to year, which will be very important when long-term trends are investigated. For example, the East Hartford, CT PAMS site did not report hydrocarbon data during the month of July 1995 because of instrumentation problems.

3.2 SITE SELECTION

Table 3-1 provides a list of the 1-hr and 3-hr data that were investigated for use in this analysis. All sites considered were PAMS type II except Harrington Beach, WI which is a PAMS type III site. This site was considered because of the paucity of data from sites that did not implement RFG. The following issues were considered in selecting data for downloading from AIRS, validation, and use in analysis:

- Data Coverage. Sufficient spatial, diurnal, and temporal coverage of the data, including before and after the introduction of RFG, is required. For the best assessment of changes in ambient species composition due to RFG implementation, data were required both before and after the change in fuel. Since most fuel changes occurred in early 1995, data for 1994 were imperative for this analysis. In some cases, 1993 data were available and thus, data from two years before and two years after the

Table 3-1. Data investigated for this analysis. All sites implemented RFG during 1995 except Harrington Beach, Atlanta, Baton Rouge, and El Paso.

Site	AIRS Code	1993			1994			1995			1996		
		June	July	Aug	June	July	Aug	June	July	Aug	June	July	Aug
Bronx, NY	360050083					3[a]	3[a]	1	1	1	1	NA	NA
Chicopee, MA	250130008					1	1	1	1	1	1	1	1
E. Hartford, CT	090031003				1	1		1	I	1			
Lynn, MA	250092006	1[c]			1	1	1	1	1	1	1	1	1
Philadelphia, PA	421010004	3	3	3	3	3	3	3	3	3	3	3	3
Chicago, IL	170310072							3	3	3	1	1	1
Chicago, IL	170310039				3[a]	3[a]	3[a]						
Clinton Drive, TX	482011035	1	1	1	I	I	I	1	1	1	I	I	1[b]
Milwaukee, WI	550790041				3	3	3	3	3	3	3	3	3
McMillan Reservoir, DC	110010043					1	1	1	1	1	1	1	1
E. Providence, RI	440071010				3	3	3	3	3	3	3	3	3
Los Angeles, CA	060371103					3	3		3	3		3	3
El Paso, TX	481410044							1	1		1	1	1
Harrington Beach, WI	550890009				3	3	3	3	3	3	3	3	
Atlanta, GA	130890002							3			3	3	3
Baton Rouge, LA	220330009	3	3	3	3	3	3	3	3	3	3	3	3

Shading highlights the available data.

Blank cells indicate no data were collected.

I = Data invalidated by the reporting agency.

NA = Data not yet available in AIRS.

Numbers 1 and 3 indicate the sampling duration (i.e., 1-hr, 3-hr).

[a] Total NMHC not reported; thus, unidentified not calculated.

[b] Data only available for August 20-31.

[c] Data only available for June 3-18.

change in fuel could be assessed. Data from both RFG and non-RFG sites were sought for comparison purposes. While data were available from RFG sites (11 sites investigated here), unfortunately, data from non-RFG sites in El Paso, TX and Atlanta, GA were not available prior to 1995.

- Measurement techniques. Different monitoring quality assurance procedures may significantly influence data availability and data quality. For example, use of a Nafion dryer with the auto-GC systems may effect the unidentified mass (typically lower with auto-GC systems), concentrations of some species, and the ability of the system to report some species compared to the use of canisters. Problems have been noted with the use of a Nafion dryer for i-butene (whose concentration may change significantly with RFG implementation). Note in Table 3-1 that at the Bronx and Chicago, the measurement systems changed from 3-hr canisters in 1994 to 1-hr auto-GCs in 1995; this change in sampling could have an effect on our ability to discern a trend solely due to fuel changes.

- Quality of ambient data. It is critical to evaluate data quality and perform additional data validation, if necessary, before performing analyses, in order to ensure the analysis results are driven by robust data and not by invalid data. In addition, the influence of monitoring methods and procedures on data accuracy and precision, and thus on analysis results, should be understood and taken into account. Data validation for this project is discussed in Section 3.4.

- Specific species measured in the PAMS or other programs. Note that some important species (e.g., MTBE, i-butene, and all alcohols) are not currently measured in the PAMS program. In addition, although several of the carbonyl species may be useful in evaluating RFG, the PAMS carbonyl measurements are less accurate and less precise than speciated hydrocarbon measurements and were often not available at many sites.

- Data completeness at the sites from year to year. Changes in data completeness can also confound trend analysis. For example, while data were available at the Clinton Drive (Houston), TX PAMS site in 1993 and 1995, data were not available in 1995 and 1996 due to analytical problems (data were invalidated by the reporting agency and thus, unrecoverable). More specifics regarding this issue are discussed later in this section.

The data shown in Table 3-1 that were not already in STI's possession were retrieved from AIRS. A preliminary assessment of data validity and completeness was performed including contact with the reporting agencies in Georgia and Texas in hopes of obtaining additional data prior to 1995. The preliminary assessments indicated that data were available from two non-RFG sites (Baton Rouge, LA; Harrington Beach, WI) and from 8 RFG sites (Springfield, MA - Chicopee; E. Hartford, CT; Boston, MA - Lynn; Philadelphia, PA; Milwaukee, WI; Washington, D.C. - McMillan Reservoir; E. Providence, RI; and Los Angeles, CA) that did not have significant location (Chicago) or analytical changes (Chicago, Bronx).

3.3 SAMPLE SELECTION

There are several factors involved in selecting the data to be used in an analysis of ambient species trends. First of all, the focus of this investigation is on data collected during the ozone season when most of the hydrocarbon samples are collected. While the ozone season may vary among the various regions, with some regions having to collect data during a longer time period than others, nearly all regions collect hydrocarbon data from June through August. Thus, this analysis focused on data collected during these three summer months. The following issues need to be considered in this selection:

Potential influence of reactivity

The loss of material by photochemical reaction, including the relative loss of one species versus another due to differences in reactivity, can significantly influence the trends in species or ratios. One way to minimize the effect of reactivity is to select data that were collected when emission rates were high and chemical reaction rates were low. Most analyses of this type have focused on the early morning traffic rush of 0600 to 0900 local time. The 3-hr PAMS samples are supposed to collect data during this time period even when only three samples are collected a day. Thus, while data for the entire day were retrieved and validated, most analyses were carried out using the morning rush hour data.

Importance of transport and background concentrations

Several issues related to the transport of pollutants from distant areas may influence ambient concentrations. These issues can be addressed by assessing the predominant meteorology at the PAMS sites, including wind direction and wind speed before and during the hours selected for data analysis, and the air parcel age using hydrocarbon species ratios and ozone and NO_y concentrations. Where possible in this investigation, meteorological and other air quality (e.g., ozone) data were obtained. However, these data were not available for all sites and for all samples. During the early morning hours, we have found that the concentrations are sufficiently high to indicate that fresh emissions were dominant at the site.

Potential influence of meteorology, and other parameters, on ambient concentrations

The diurnal variation and daily peaks in temperature and mixing height may have a significant influence on hydrocarbon concentrations and sample composition. Methods to investigate and/or adjust for variations in meteorology and other parameters which might influence ambient concentrations need to be considered.

Other influences

Data were identified by day of week, begin hour, month, and year for use in generating statistics. Preliminary investigations of the results segregated by weekend versus weekday showed no significant difference in this analysis. Since the largest number of samples needed to be retained as possible, all data were used in the analyses.

3.4 DATA VALIDATION

The issue of data validation was important for this project because of the use of ambient data in the analysis. Data validation is needed because serious errors in data analysis results can be caused by erroneous individual data values. As stated by the EPA: "The purpose of data validation is to detect and then verify any data values that may not represent actual air quality conditions at the sampling station" (U.S. EPA, 1984, Sec. 2.0.3, p.10). Data validation and quality control is performed by the reporting agencies prior to submittal of the data to AIRS. However, our experience with these data has shown that additional inspection of data is imperative as many data problems have not been removed prior to data submittal. This is particularly crucial for 1993 and 1994 data sets.

We define data validation as the process of determining the quality of the observations and identifying their validity. The data validation process is important because it helps identify data with errors, biases, or physically unrealistic values (which may include outliers) <u>before</u> they are used for analyses or modeling. The objectives of the data validation process are to produce a database with values that are of a known quality; to evaluate the internal, spatial, temporal, and physical consistency of the data; and to intercompare data to identify errors, biases, or outliers. Outliers are defined as data values that are physically, spatially, or temporally inconsistent with the vast majority of observed values. The identification of outliers is particularly important and particularly difficult in the analysis of speciated NMOC data, as the concentration distributions of these species are typically broad and some extreme values may be the result of measurement error. Data validation is an ongoing process, and each new analysis provides an additional opportunity to gain confidence in the validity of the data.

As a part of the NARSTO-Northeast study, STI developed a Windows-based, menu-driven program which provides a graphical platform from which to display the VOC data collected with automated GC systems as a part of the PAMS network, to perform quality control tasks on the data, and to begin data analysis (Main et al., 1996a). This software was used to validate and initially explore the hydrocarbon and carbonyl data in this project. The 1-hr or 3-hr concentrations of the PAMS target species were displayed in three ways: scatter plots, fingerprints, and time series. Using scatter plots, the relationship between species at one site or at a number of sites may be investigated. Fingerprint plots show the concentration of each species in a sample (in roughly chromatographic order) and help to identify unique characteristics of the samples. Time series plots show the concentrations of species in every sample over a specified time period and are useful in showing the diurnal behavior of a species. The QA software also computes some species groups including total unidentified hydrocarbons, total aromatic hydrocarbons, total olefins, total paraffins, and total carbonyl compounds. The weight percent of the total NMHC of individual species and species group totals may be computed and displayed in the same manner as the concentration data.

This software was implemented to review PAMS VOC data; the software facilitates the inspection of diurnal patterns, concentration spikes, the periodicity of concentration increases or decreases, the relationship between various species, and hour-by-hour inspections of the sample fingerprints. These investigations assist the experienced analyst in rapidly establishing site-specific patterns in concentration, determining specific hours for which there are questions, and

discovering problems in the data set which need correction. For example, one site-month of hourly VOC data contains approximately 720 samples with 60 individual hydrocarbon and carbonyl species, four species groups (i.e., olefins, paraffins, and aromatic and unidentified hydrocarbons), and the total NMHC for a total of about 47,000 data entries; this data set can be relatively thoroughly investigated in about 2 hours.

Data validation had been performed on much of the data listed in Table 3-1 as a part of separate contracts. As a part of a project for the Coordinating Research Council, STI had previously obtained and validated VOC data from 21 PAMS sites in the Northeast for the summer of 1995. As a part of the PAMS data analysis workshops conducted for the EPA OAQPS (e.g., Main et al., 1997a), STI also obtained and validated VOC data collected in 1994, 1995, and/or 1996 from many areas in the United States. The data not investigated in the previous projects were validated in a similar manner as discussed above. The results of data validation investigations for the current study are summarized here. Of the ten cities with RFG that were investigated, five cities have data that should be adequate for the analyses (Springfield, MA - Chicopee; Boston, MA - Lynn; Philadelphia, PA; Milwaukee, WI; Washington, D.C. - McMillan Reservoir), two cities have data that may be adequate depending upon assessment of data gaps (E. Hartford, CT; Los Angeles, CA), and three cities have data that are problematic (Bronx, NY and Chicago, IL experienced analytical changes between 1994 and 1995; Chicago, IL also experienced a significant site change during this time period; data are unavailable at Houston, TX - Clinton Drive - for 1994 and much of 1996). Of the five cities without RFG implementation that were investigated, only two cities have data that may be adequate depending upon further investigation of the data quality (Baton Rouge - Capitol, LA) and data quantity (Harrington Beach, WI). Data were not available for the 1994-1996 time period at Pittsburgh, PA; Atlanta, GA; and El Paso, TX.

Specific data validation findings are summarized in Appendix A. The tables in the appendix list the species that were flagged as suspect or invalid and provide reasons why the data were flagged. For regions that implemented RFG in 1995, the following summarizes the data availability and comments regarding the data validity (note that flagged samples, both suspect and invalid, were not included in subsequent analyses):

- Lynn, MA (Tables A-1 through A-4). Data from the summers of 1993-1996 were obtained. The 1994-1996 data were generally of very good quality. There were only two weeks of data available in 1993; there were problems with the NMHC values for these data and misidentification problems with 2- and 3-methylpentanes. In 1994, there were several samples with missing C2-C6 data (i.e., one of the columns failed), calibration gas carryover affecting the heavier species, and misidentification problems. The 1995 and 1996 data showed few problems.

- Philadelphia, PA (Tables A-5 through A-8). Data from June-August for 1993-1996 were obtained and the data were generally of fair quality. The 1993 data exhibited problems affecting some samples including high concentrations of 3-methyl-1-butene and misidentification of cyclopentane, hexane, and 2,3-dimethylbutane. Fortunately, most of these problems appeared to occur during the midday and thus, did not cause elimination of very many early morning samples. In 1994, there were still many samples with high concentrations of 3-methyl-1-butene and additional problems with t-2-pentene and 2-

methyl-2-butene. In 1995, the o-xylene data are all invalid (detection limit too high to be useful). In 1996, there were few problems evident in the data.

- Chicopee, MA (Tables A-9 and A-10). Data from June-August for 1994-1996 were available and data quality was very good. In 1994, there were samples with acetylene concentration problems, xylenes misidentification, and high n-undecane concentrations due to calibration gas carryover. In 1995-1996, some data showed evidence of misidentification of 2-methylpentane and calibration gas carryover affecting the higher molecular weight species.

- E. Hartford, CT (Tables A-11 and A-12). Data from 1994 and 1995 were available; however, August data were invalid for 1994, and July data were invalid for 1995. Data were not collected at this site in 1996. Data quality in general was good (there was too much missing data during 1994-1995 to warrant a label of "very good").

- Houston, TX (Clinton Drive) (Table A-13). The data from 1993-1995 were obtained and data were of good quality. The 1993 data were not available in AIRS; these data were obtained from STI's Coastal Oxidant Assessment for Southeast Texas (COAST) database. Through discussions with the Texas Natural Resource Conservation Commission (TNRCC), we learned that the data were invalidated by the reporting agency for June through August 1995 and June through most of August 1996. In addition, the 1995 and 1996 data did not report the NMHC values. Thus, the Clinton Drive data could not be used in assessing trends because of the large data gaps.

- Milwaukee, WI. Data were obtained from the summers of 1994-1996 and were generally of very good quality. No data warranted changes in QC status, however, in several samples, 2-methyl-2-butene concentrations were somewhat higher than observed at other sites. Also, the June 4 through 13, 1994 data reported a different speciation list than June 14 through August 31, 1994.

- Washington, D.C. Data from the summers of 1994-1996 were obtained and were of very good quality. Only one sample was invalidated (reported zero concentrations for all species) and no samples were flagged as suspect.

- Los Angeles, CA. Data were obtained for the summers of 1994-1996 for the site at North Main. The 3-hr samples at this site were collected every third day during July and August. The data were of good quality (i.e., no discernible problems) and no samples were flagged.

- Bronx, NY (Table A-14). Data were obtained from July 1994 through June 1996 and were generally of good quality. However, the sampling period and analytical methods were changed between 1994 and 1995 (from 3-hr canisters every third day to 1-hr auto-GC on a daily basis) and this change may influence comparisons between 1994 and 1995. In the 1996 data, 2,3-dimethylbutane appears to have misidentification problems; in 1994, no NMHC was reported. These data are probably not suitable for assessing 1994-1995 trends.

- E. Providence, RI (Table A-15). Data were obtained for the summers of 1994-1996. The 3-hr samples at this site were collected daily and were generally of good quality. During 1994, many of the samples were missing C2/C3 data and, thus, the total NMHC values

provided were invalid. During 1995, a few samples were set to suspect because of high concentrations, a suspected C2/C3 column problem (ethane concentration of zero when higher carbon number species were greater than 1 ppbC), and a calibration gas sample. In 1996, three samples had C2/C3 problems.

- Chicago, IL. Data were obtained for the summers of 1994-1996. The 1994 and 1995 data were 3-hr, collected every third day. The 1996 data were 1-hr data collected daily. Data were of fair quality. NMHC data were not reported in 1994. There was a significant site change between 1994 and 1995 (from Northwestern University to a site on the lake shore). In 1996, the 1,2,4-trimethylbenzene data were suspect all summer; the detection limit appeared to be about 4.5 ppbC. There were also several o-xylene outliers that may indicate an independent source of this species from the other two isomers.

For regions that have not implemented RFG:

- Baton Rouge, LA (Tables A-16 and A-17). Data from 1993-1996 were obtained. The 1993-1995 data were collected on a 3-hr, daily basis; the 1996 data were collected every third day. Data were of poor to fair quality and analysis results should be viewed qualitatively at best. STI evaluated the data set and later reviewed Stoeckenius et al., 1996 which discussed the Capitol data in some detail. Many of the 1993 samples have invalid NMHC values (i.e., the sum of species concentrations exceeded the reported NMHC). Throughout the data set (all years) the acetylene concentrations appeared low and often zero which may indicate a problem. Stoeckenius et al. also noted the acetylene problems and the acetylene data should be considered questionable. In addition, Stoeckenius et al. noted that ethane and ethene may not have been well-resolved in 1994 due to high humidity effects on GC column retention times. (Presumably, this would also apply to 1993 data as well.) The isopentane in 1994 appears to have been incorrectly identified as 3-methyl-1-butene in the 1994 data; however, the 1993, 1995, and 1996 data did not show this problem. There were also other problems in the data set of misidentification, missing data, and abnormally high concentrations. For example, the high concentrations of isobutane in 1994 were possibly due to coelution with methanol (Stoeckenius et al., 1996). The species identification of 2-methylpentane and 3-methylpentane appeared troublesome in both 1993 and 1994.

- Harrington Beach, WI (Table A-18). Data were obtained from 1994-1996 and were generally of good quality. These data were collected in 3-hr canisters on every third day and during high ozone episodes. Typically, only three samples were collected during the day. For many samples, the NMHC values were invalid (i.e., reported only as the sum of the identified). In general, toluene concentrations appeared low and methylcyclopentane concentrations appeared high compared to other cities. This site is downwind of Milwaukee.

- Other sites: No data were available in AIRS or from the reporting agencies for Atlanta or El Paso prior to 1995.

3.5 DATA CAVEATS

Fairly extensive data validation is important when investigating the PAMS data from AIRS. In the process of performing data validation as described in Section 3.3, conversations with the reporting agencies were initiated in order to obtain possible reasons that some data appeared anomalous. This section summarizes the special attention required for each of the data sets when interpreting the results of the trends analyses. Available valid data are summarized in **Table 3-2**. To obtain the percent data available, the number of valid samples was divided by the number of samples planned to be collected. In addition to the routine sample collection, at some sites extra samples were collected on an episodic basis (i.e., when ozone concentrations were forecast to be high). Thus, these sites show more than 100 percent data available. Note that for all data submitted to AIRS, the sample time should be reported as standard time. Thus, for 0600-0900 local time in the summer, the samples should be 0500-0800 ST. The data submitted to AIRS however, showed that sometimes, samples appeared to be reported in local (daylight) time (e.g., three hour samples starting at 0600). To accommodate the possibility of time convention problems, the 0500, 0600, 0700, and 0800 data were retained from the hourly files and the 0500-0800 and 0600-0900 data were retained from the 3-hr files.

Table 3-2. Available number of valid samples (n) collected during the 0500 to 0900 time period and the percent data completeness for June through August of each year (sampling times are provided in Table 3-1).

Sites	1993		1994		1995		1996	
	n	%	n	%	n	%	n	%
Lynn, MA	24	7	264	72	344	93	318	86
McMillan R., D.C.			120	37	328	89	220	60
Chicopee, MA			264	72	233	63	298	80
Philadelphia, PA	41	45	75	82	69	75	83	90
Milwaukee, WI			34	113	35	117	68	113
E. Providence, RI			34	37	80	87	81	88
Los Angeles, CA[a]			13	43	24	80	21	70
Bronx, NY			28	30	269	73	50	14[b]
E. Hartford, CT			120	33	127	35	0	0
Chicago, IL			34	117	36	120	354	96
Baton Rouge, LA	72	78	82	88	80	87	31[c]	100
Harrington B., WI			33	110	33	110	20	65

[a] Only July and August data were reported.
[b] Only June data was available in AIRS at this time.
[c] Changed from daily to periodic sampling.

The following data caveats should be considered when evaluating trends in possible RFG indicators:

- Lynn, MA. There is little wrong with this data set. Only seven days of valid data are represented in 1993; thus, the sample size is too small to be quantitative for this year. The 1993 data have been included for qualitative use only.

- Washington, D.C. There are no overwhelming problems with this data set and it is well-suited for a trend analysis.

- Chicopee, MA. The 1994 acetylene and xylenes data showed problems and analyses using these species should not be over-interpreted.

- Philadelphia, PA. In 1995, the o-xylene data are all invalid (detection limit too high to be useful) and, thus, analyses using the xylenes should not be performed.

- Milwaukee, WI. The available valid data accounted for over 100 percent of the possible samples in all three years (every third day plus some ozone episodes). However, the number of samples per month is small and may limit the percent change in ambient concentrations that can be statistically observed.

- E. Providence, RI. The missing C2/C3 data in 1994 for many samples means that the weight percent and ratios with acetylene will not be as useful. There were few problems with the 1995 and 1996 data.

- Los Angeles, CA. The number of samples per month is small and may limit the percent change in ambient concentrations that can be statistically observed.

- Bronx, NY. These data are probably not suitable for assessing 1994-1995 trends. First, only 30 percent of the data in 1994 are represented. Second, because the sampling period and analytical methods were changed between 1994 and 1995 (from 3-hr canisters every third day to 1-hr auto-GC on a daily basis), trends between these two years due to changes in ambient species concentrations may be difficult to prove. Finally, the NMHC values were not reported for 1994, thus weight percents cannot be calculated.

- E. Hartford, CT. The principal problem with the data at this site is the lack of data during 1994-1996. Because of the missing data, the available valid data for this site are low, only about 35 percent each in 1994 and 1995, and data were not collected in 1996. Because of the paucity of data at this site, the data are less useful for trend analysis and should probably be assessed only qualitatively.

- Chicago, IL. There was a significant site change between 1994 and 1995 (from Northwestern University Medical School in downtown Chicago to a site on the lake shore) and, thus, any trends between these two years should be assessed carefully. The 1,2,4-trimethylbenzene data should not be used because of the apparent high detection limit.

For regions that have not implemented RFG:

- Baton Rouge, LA. Data at this site in 1993 and 1994, particularly, are of only fair quality and analysis results should be viewed qualitatively at best. Problems with the NMHC values, acetylene, isobutane, and isopentane, render many of the ratios and weight percent indicators nearly useless for analysis at this site for 1993 and 1994.

- Harrington Beach, WI. The available valid data accounted for 100 percent of the possible samples in 1994 and 1995 (every third day plus ozone episodes). This site is a PAMS type III (all other sites are PAMS type II). The site is located downwind of Milwaukee and, thus, ambient concentrations may be influenced by transport of hydrocarbons from Milwaukee. One might expect mixed results when investigating changes at this site because locally, RFG was not required, while in the predominant direction upwind, RFG was used.

In summary, there are sufficient valid data over the 1994-1996 time period at sites that implemented RFG for further analysis. The picture is less rosy for sites that did not implement RFG. The Baton Rouge data are questionable during 1993 and 1994 and Harrington Beach data may be influenced by Milwaukee upwind.

4. INVESTIGATION OF RFG EFFECTS ON AMBIENT VOC CONCENTRATIONS

4.1 DATA PREPARATION AND INDICATORS

In order to assess trends in the ambient hydrocarbon concentrations hypothesized to occur with RFG, the validated PAMS data were first merged into SPSS:Systat, a statistical software program, to prepare files for each site that contained early morning, summer data from each year. Once the database was prepared, the statistics described in Section 2 could be performed. Box plots were used to visualize the data. As an example, **Figure 4-1** shows an annotated box-whisker plot. The box shows the 25th, 50th (median), and 75th percentiles. The whiskers always end on a data point, so when the plots show no data points beyond the end of a whisker, the whisker shows the value of the highest or lowest data point. The whiskers have a maximum length equal to 1.5 times the length of the box (the interquartile range, IR). If there are data outside this range, the "outliers" are also further identified with asterisks representing the points that fall within three times the interquartile range and circles representing points beyond this.
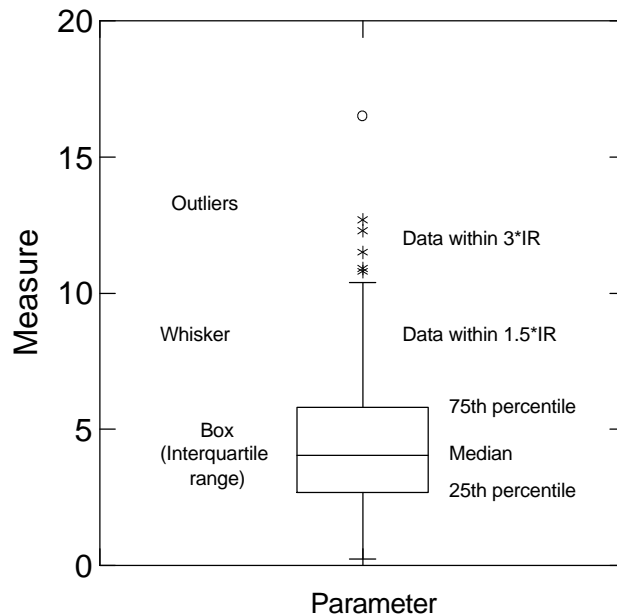
Figure 4-1.   Annotated box-whisker plot with outliers.

Box plots were prepared for several species concentrations, weight fractions, and ratios as listed in **Table 4-1**. Unfortunately, not all of the indicators discussed in Section 2.2 are available in the PAMS data sets, and some species are not available for all years. For example, the species i-butene and 1,3-butadiene are not measured routinely at many PAMS sites because of low concentrations and possible analytical problems associated with drying

Table 4-1.   Species (concentration and weight percent of NMHC) and ratios used as possible indicators of RFG implementation.

Indicator	Motivation
Benzene	Specific reductions called for by RFG regulations.
n-Butane	Normally reduced to lower RVP; RVP needs to be reduced for RFG.
1,3-Butadiene	Modeling studies predict a 20 percent decrease.
Xylenes	An aromatic compound; aromatics need to be reduced for RFG.
Formaldehyde	Modeling studies predict a 20 percent increase.
i-Butene	Decomposition product of MTBE; MTBE is an oxygenate used to meet RFG requirements.
Trimethylbenzenes (TMB)	An aromatic compound; aromatics need to be reduced for RFG, particularly C9-C10 aromatics.
C8 Alkanes	Possibly substituted for aromatics to meet RVP reductions.
NMHC	RFG should result in a reduction in the total amount of hydrocarbons.
Benzene/Toluene	If NMHC is not available, this is a good surrogate for the benzene weight percent since toluene is typically measured accurately and does not change as much with RFG as benzene.
n-Butane/i-Pentane	If NMHC is not available, this is a good surrogate for the n-butane weigh percent since i-pentane is not expected to change with RFG.
Benzene/Acetylene	If NMHC is not available, this is a good surrogate for the benzene weight percent since acetylene does not change with RFG and has few sources other than motor vehicle exhaust.
TMB/Toluene	If NMHC is not available, this is a good surrogate for the TMB weight percent since toluene is typically measured accurately and does not change as much with RFG as TMB.
TMB/Xylenes	TMB expected to change more than total xylenes.

during sample collection. Estimates by Ligocki and Yarwood (1994) showed that C9-C10 aromatics might be a possible indicator of RFG changes; however, the list of C9-C10 aromatics changed from year to year at the PAMS sites. Therefore, at many sites, it was not possible to formulate a sum of C9-C10 aromatics that was consistent over all years. Instead, an analysis of the most abundant of these species, 1,2,4-trimethylbenzene, was performed. Similarly, for the C8 alkanes, 2-methylheptane, n-octane, and 3-methylheptane were investigated separately. Formaldehyde data were sparse both temporally and spatially.

To compare the annual data statistically, a common method is the two-sample t-test that compares the means of the two data sets such as the data for 1994 versus 1995. However, since the ambient hydrocarbon data can often have a very skewed distribution (e.g., lognormal), a comparison of means may not be the most appropriate. Nonparametric tests do not assume that the data conform to a particular probability distribution. Nonparametric tests compute nonparametric statistics for groups of cases. Nonparametric models are often appropriate when the usual parameters, such as mean and standard deviation based on normal theory, do not apply. A Kruskal-Wallis analysis of variance (ANOVA) on the selected variables was used. The Kruskal-Wallis test and Mann-Whitney U-statistics are analogous to the independent groups t-test. Data sets were considered statistically different when the p-values were less than 0.05.

4.2 RESULTS

4.2.1 Overview

In this project, the trends in more than 20 species concentrations, weight percents, or species ratios were investigated for every site. This investigation consisted of a review of annual box plots of the data and a statistical comparison of the annual data. Both nonparametric tests and t-tests were applied to the annual data. In most cases, the results were the same. For example, when the Kruskal-Wallis ANOVA indicated p-values less than 0.05, the t-test results also indicated p-values less than 0.05. **Tables 4-2 and 4-3** show a summary of the results from these tests on the 1994/1995 and 1995/1996 data sets. In these tables, "+" indicates that the indicator showed a statistically significant increase (p-value<0.05 based on the Kruskal-Wallis test) from 1994 to 1995 or 1995 to 1996, "-" indicates that the indicator showed a statistically significant decrease, and a blank indicates that the indicator did not show a statistically significant change. The actual p-values for these summary tables are provided in Appendix B. In a few cases, the indicator showed a statistically significant change, or the results from the t-test and nonparametric test disagreed. If the p-values from one test were <0.05 but not from the other, both results are indicated in the table. Also listed in the tables is "na", data inadequate for analysis (e.g., missing NMHC or indicator for one of the two years, species not analyzed for, etc.).

For the Bronx and Chicago data, the sampling intervals shifted from 3-hr to 1-hr during the 1994-1996 time period. Thus, more samples were available for the years with the 1-hr data than with the 3-hr data. Statistics were computed in two different ways for these data sets:

- The 1-hr samples collected were compared one hour at a time to the 3-hr samples in the previous year. For example, the 1-hr samples collected at 0600-0700 in 1995 were compared to the 3-hr samples collected during 0600-0900 in 1994.

- All the 1-hr samples collected were compared to the 3-hr samples in the previous year. For example, the total number of 1-hr samples collected at 0600-0700, 0700-0800, and 0800-0900 in 1995 were compared to the 3-hr samples collected at 0600-0900 in 1994.

Note that the 1-hr samples were not averaged across the 3-hr sampling period. We did not want to "dilute" the greater detail and larger pool of available data provided by the 1-hr samples. The results of these two approaches were compared. In most cases, the general trend of the data was the same using either one hour from the 1-hr samples (e.g., 0600-0700) or all the 1-hr samples (e.g., the 0600, 0700, and 0800 samples). The results in the tables reflect the p-values obtained using all the 1-hr samples.

Table 4-2. Statistically significant changes from 1994 to 1995 in species concentrations, weight fractions, and ratios. Increases between 1994 and 1995 are signified with a "+", decreases with a "-", and no change with a blank. Statistical significance was

determined using the Kruskal-Wallis nonparametric test, differences with p-values less than 0.05 were deemed significant. When the nonparametric test and t-test results disagreed, the parameter before the "/" is the nonparametric result while the value after the "/" is the t-test result. P-values are provided in Table B-1.

Species, Ratios	Los Angeles, CA	Chicago, IL	New York City (Bronx)	Springfield MA (Chicopee)	E. Hartford, CT	Boston, MA (Lynn)	Wasahington DC (McMillan Reservoir)	Philadelphia, PA	Milwaukee, WI	E. Providence, RI	Harrington Beach, WI	Baton Rouge, LA
Concentrations:												
n-Butane		-		-		-	-	-	-	"-/"	-	
1,3-Butadience		na	na	na	na	na	na	na	na	na	na	na
Benzene		-	-	-	-	-	-	-	-		-	"/+"
Xylenes		-		-	"-/"	-	-	-	-		-	
Formaldehyde	na		na	"+/"	"-/"	-	na		na		na	na
i-Butene	+	na	na	na	na	na	na	na	na	na	na	na
Trimethylbenzenes		-		-	"-/"	-	-	-	-		-	
TNMOC	+	na	na		"+/"	+		-	-		-	+
Isoprene	+	"/-"		+			"/+"		"/-"			"+/"
Weight Percent:												
n-Butane	-	na	na	+		-	-	-	-	-		-
1,3-Butadience		na	na	na	na	na	na	na	na	na	na	na
Benzene	-	na	na	-		-	-	-	-	"-/"	"-/"	-
Xylenes	-	na	na	-			-					-
i-Butene	+	na	na	na	na	na	na	na	na	na	na	na
Trimethylbenzenes	"/-"	na	na	-	"-/"		-					"/-"
Isoprene		na	na	+	+			+				
Ratios:												
Benzene/Toluene	-	+	-	+	+	-	-	-		"-/"		+
n-Butane/i-Pentane	-	"-/"	+	"/-"	-	-	-	-	-	-	-	-
Benzene/Acetylene	-	+	"+/"		+	-	-	-	-	"-/"	-	+
TMB/Toluene		+		+			-	-				
TMB/Xylenes		+	+	+			-	-		+		

na = data inadequate for the analysis (e.g., missing NMHC, not analyzed for, etc.)

RFG was not implemented at Harrington Beach or Baton Rouge.

Table 4-3. Statistically significant changes from 1995 to 1996 in species concentrations, weight fractions, and ratios. Increases between 1995 and 1996 are signified with a "+", decreases with a "-", and no change with a blank. Statistical significance was determined using the Kruskal-Wallis nonparametric test, differences with p-values less than 0.05 were deemed significant. When the nonparametric test and t-test results disagreed, the parameter before the "/" is the nonparametric result while the value after the "/" is the t-test result. P-values are provided in Table B-1.

Species, Ratios	Los Angeles, CA	Chicago, IL	New York City (Bronx)	Springfield MA (Chicopee)	E. Hartford, CT	Boston, MA (Lynn)	Wasahington DC (McMillan Reservoir)	Philadelphia, PA	Milwaukee, WI	E. Providence, RI	Harrington Beach, WI	Baton Rouge, LA
Concentrations:												
n-Butane			-		n		+		+	-		+
1,3-Butadience		na	na	na	o	na	na	na		na		na
Benzene	-	-	-	"/-"			"+/"	"+/"	+	-		
Xylenes		-	"/+"	"/-"	d	-	+		+	-		
Formaldehyde	na	-		"-/"	a		-		-	-	-	-
i-Butene		na	na	na	t	na	na	na	na	na	na	na
Trimethylbenzene		-		+	a	-	+	+	+	-		-
TNMOC		-	+	-		-	"/+"			-	-	-
Isoprene	"+/"					-	+	"-/"	+	-		
Weight Percent:												
n-Butane		+	-	+		+	+	"-/"			+	+
1,3-Butadience		na	na	na		na	na	na		na	na	na
Benzene	-	-	-	+		+	"/+"		+	-	+	+
Xylenes				+			+	-	"+/"	-		+
i-Butene		na	na	na		na	na	na	na	na	na	na
Trimethylbenzene	"/+"	-		+			+			-		
Isoprene		-		"/+"			+	"-/"		"-/"		+
Ratios:												
Benzene/Toluene	-	-	"-/"	"+/"		+			+	"+/"		
n-Butane/i-Pentan	-	+	-	+			+	-	+			"-/"
Benzene/Acetylen	-	-	-	"-/"		"+/"	-		+			"-/"
TMB/Toluene		+	-	"/-"		-	+	+	-	-		-
TMB/Xylenes		"/+"	-	-		"+/"	+	+	-	-		-

na = data inadequate for the analysis (e.g., missing NMHC, not analyzed for, etc.)

RFG was not implemented at Harrington Beach or Baton Rouge.

In general, the benzene weight percent and other benzene indicators (e.g., benzene/toluene ratio) showed statistically significant declines from 1994 to 1995. From 1995 to 1996, there were

4-5

fewer statistically significant changes. Since much of the focus of RFG has been on the apparent ambient benzene reductions, **Table 4-4** summarizes the change in the ambient benzene weight percent between 1994 to 1995 and 1995 to 1996. Significant decreases in summer morning ambient benzene weight percents were observed at most sites between 1994 and 1995 (data were not available at all sites), ranging from about -0.4 to -1.3 percent difference in the medians. Some sites showed continued declines in ambient benzene levels in 1996 while other sites showed increases that offset some or all of the decreases between 1994 and 1995. The results in these tables are discussed in more detail in the following sections.

Table 4-4. Absolute changes in median ambient benzene weight percent between 1994 to 1995 and 1995 to 1996 (summer morning data). Only statistically significant changes are listed (i.e., p-values < 0.05 using the Kruskal-Wallis nonparametric test). Decreases are shaded. The 1994 median weight percent of benzene is provided as a point of reference.

Site	Median wt% Benzene 1994	Absolute Change 1994-1995	Absolute Change 1995-1996
Boston (Lynn), MA	2.53	-1.1	0.44
Washington (McMillan Reservoir), D.C.	3.49	-1.3	
E. Providence, RI	2.16	-0.38	-0.18
Milwaukee, WI	3.23	-0.40	0.40
Philadelphia, PA	3.06	-0.70	
Springfield (Chicopee), MA	1.99	-0.40	0.40
Los Angeles, CA	3.08	-0.80	-0.40[a]
Chicago, IL	2.89[b]	ns	-0.70
New York City (Bronx), NY	1.93[b]	ns	-0.70
Baton Rouge, LA[c]	2.02	-0.50	0.60
Harrington Beach, WI[c]	2.97	-0.44	1.1

[a]   NIPER fuel data show an additional decrease in fuel benzene composition between 1995 and 1996.
[b]   1995 median value.
[c]   RFG not implemented.
ns   Data not suitable (e.g., cannot compute weight percent because of missing NMHC values).

4.2.2 Results for Sites that Implemented RFG

At the following sites, ambient benzene concentrations and/or weight percent and other indicators showed significant declines. In other words, there was consensus among several indicators that RFG may have influenced the ambient concentration and composition. Specific findings are as follows:

- Lynn, MA (Boston). The benzene concentrations, weight fractions, and ratios all decreased between 1994 and 1995 (**Figure 4-2**). Statistical tests of the benzene data confirmed that the changes were statistically significant. Fuel composition data (e.g., Table 2-3) showed that fuel benzene content was reduced. Interestingly, benzene weight percent showed a statistically significant increase between 1995 and 1996. OMS

4-6

compliance survey data for the Boston area (Table 2-7) showed that fuel benzene content also increased.

Lynn, MA

Figure 4-2.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent (BENZW), (c) benzene to toluene concentration ratio (BT), and (d) benzene to acetylene concentration ratio (BACETY) at Lynn, MA for the summers of 1993-1996. Data are June-August, 0500-0800 EST, valid data only. The small amount of data available in 1993 are shown for qualitative purposes only. The statistical tests indicated that the 1994 and 1995 data were different.

- McMillan Reservoir (Washington, D.C.). The weight percent and concentrations of benzene, total xylenes, sum of trimethylbenzenes, and related ratios decreased between

1994 and 1995 (e.g., **Figure 4-3**). The statistical test results indicated that the 1994 and 1995 data were significantly different. Table 2-3 shows that the fuel benzene content declined during this time period. Increases in the fuel benzene content shown in Table 2-7 for 1995 to 1996 were not paralleled by increases in the ambient benzene. However, butane concentration and weight percent did show a statistically significant increase from 1995 to 1996.

- E. Providence, RI. Benzene and n-butane weight percents, and related ratios (e.g., benzene/toluene, n-butane/i-pentane) decreased between 1994 and 1995 (e.g., **Figure 4-4**). The statistical test results indicated that the differences between 1994 and 1995 were significant. The fuel benzene content was shown to decline in the Northeast and mid-Atlantic (see Table 2-5) during this time period. From 1995 to 1996, most of the species concentrations and weight percents showed statistically significant declines. However, the compliance survey data (Table 2-7) does not indicate additional declines in fuel benzene, for example.

- Philadelphia, PA. The weight percent and concentration of benzene and related ratios decreased between 1994 and 1995 (**Figure 4-5**). The statistical tests indicated that the 1994 and 1995 data were significantly different while the 1993/1994 and 1995/1996 data were not. These results are consistent with fuel changes during the 1994 to 1995 (Table 2-3) and 1995 to 1996 (Table 2-7) periods.

- Bronx, NY. The benzene concentrations and benzene/toluene ratio declined between 1994 and 1995 (**Figure 4-6**). The statistical tests indicated this 1994 to 1995 decrease was significant. These results are consistent with fuel changes during the 1994 to 1995 time period (Table 2-3). Additional declines in ambient benzene were observed between 1995 and 1996, however, the fuel benzene content did not appear to change (Table 2-7).

- Los Angeles, CA. While data have been collected at this site since the late 1980's, only 1994-1996 data were available in the AIRS database. There have been many documented changes in benzene levels in the fuel in California and declines in benzene have been noted in recent work (e.g., Hammond, 1996; Zielinska et al., 1997). In the data set available to this project, i-butene concentrations and weight fractions increased between 1994 and 1995; this is consistent with predictions for this hydrocarbon to increase with RFG (e.g., Ligocki and Yarwood, 1994). Decreases were observed in the weight fractions of n-butane, benzene, and xylenes, as well as the ratios of benzene/toluene, n-butane/i-pentane, and benzene/acetylene (**Figures 4-7 and 4-8**) and the statistical tests confirmed the differences between the annual means and medians. All these changes are consistent with RFG introduction. Table 2-2 shows that even though changes in California fuels were initiated in 1992, there was a change in the fuel between 1994 and 1995 that could account for these trends. There were additional declines in fuel benzene in 1996 with the introduction of California Phase II RFG; the additional decline in ambient benzene weight percent, concentration, and related ratios observed in 1996 probably reflects this fuel change.
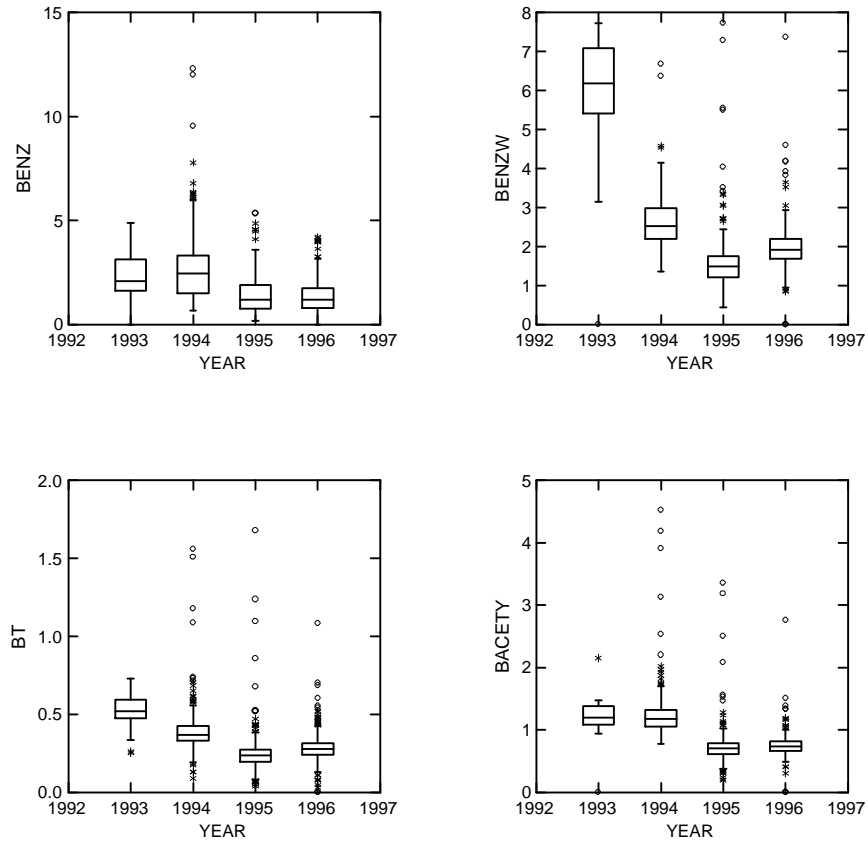
McMillan Reservoir, DC

Figure 4-3. Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent
(BENZW), (c) benzene to toluene concentration ratio (BT), and (d) benzene to
acetylene concentration ratio (BENZACETY) at McMillan Reservoir (Washington,
D.C.) for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid
data only. The statistical tests indicated that the 1994 and 1995 data were different.
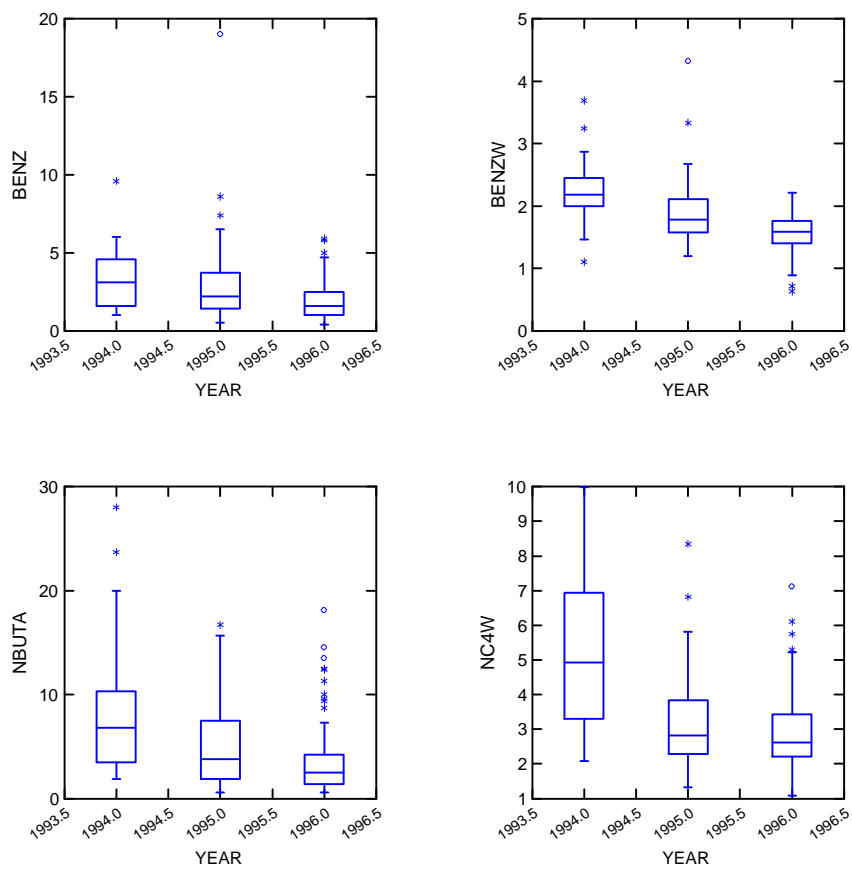
Figure 4-4.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent (BENZW), (c) n-butane concentration (NBUTA), and (d) n-butane weight percent (NC4W) at E. Providence, RI for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only. The statistical tests indicated that the 1994 and 1995 data were different.

Philadelphia, PA
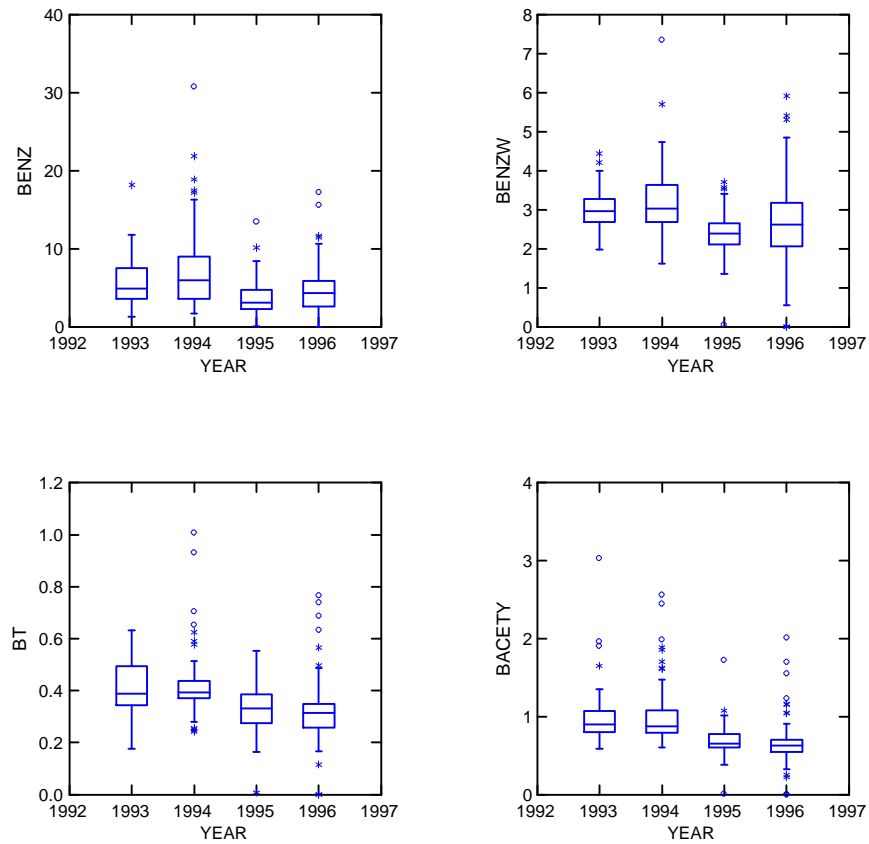
Figure 4-5.   Box plots of (a) benzene concentration, (b) benzene weight percent (BENZW), (c) benzene to toluene concentration ratio (BT), and (d) benzene to acetylene ratio (BACETY) at Philadelphia, PA for the summers of 1993-1996. Data are June-August, 0500-0800 EST, valid data only. Plots do not include three outliers so that the data are easier to interpret. The statistical tests showed that the 1994 and 1995 means were statistically different while 1993 to 1994 and 1995 to 1996 means were not.

Bronx, NY

Figure 4-6.   Box plots of (a) benzene concentration (BENZ), (b) benzene to toluene concentration ratio (BT), (c) benzene to acetylene concentration ratio (BENZACET), and (d) total xylenes concentration (TOTXYLS) at Bronx, NY for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only. Note that the 1994 data are 3-hr average while the 1995 and 1996 data are 1-hr average and only one month of data were reported for 1996 at this time.
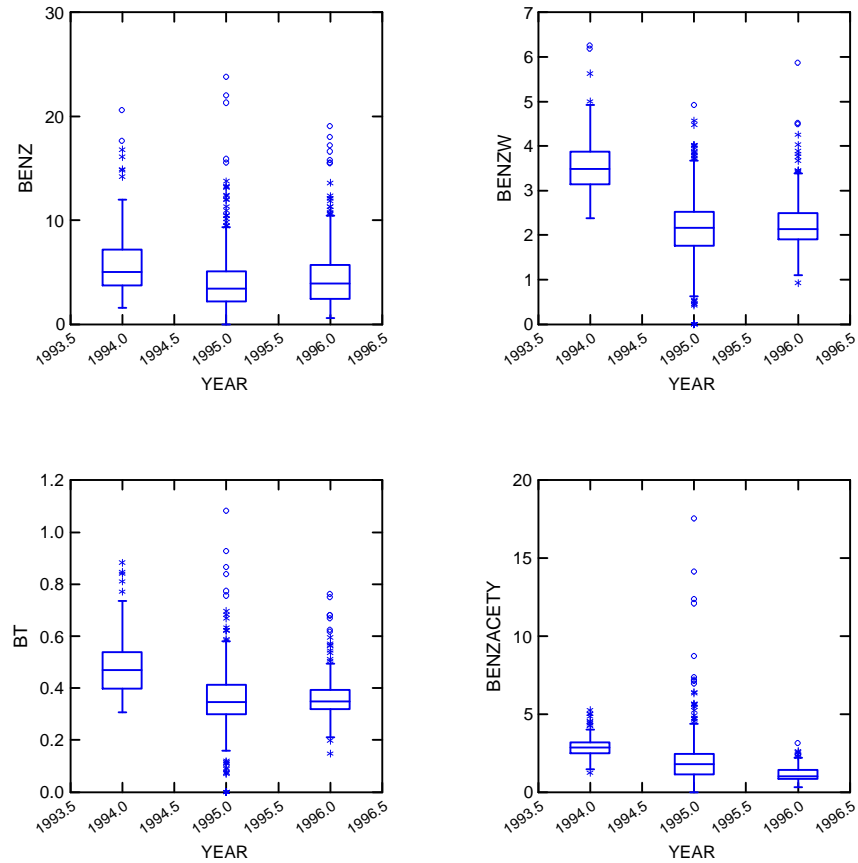
Los Angeles, North Main

Figure 4-7.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent (BENZW), (c) benzene to acetylene concentration ratio (BENZACET), and (d) benzene to toluene concentration ratio (BT) at Los Angeles, CA for the summers of 1994-1996. Data are June-August, 0500-0800 PST, valid data only. The statistical tests indicated that the 1994 and 1995 means were different for benzene weight percent and benzene ratios. Note that ambient benzene drops significantly again in 1996.

Los Angeles, CA

Figure 4-8.   Box plots of (a) 1,3-butadiene weight percent (BUTADW), (b) i-butene concentration (IBUTEW), and (c) total xylenes weight percent (XYLW) at Los Angeles, CA for the summers of 1994-1996. Data are June-August, 0500-0800 PST, valid data only. The statistical tests indicated that the means for 1994 and 1995 i-butene and the total xylenes were different.

- Milwaukee, WI. Benzene concentrations decreased between 1994 and 1995 (**Figure 4-9**). However, both the concentration and weight percent of ambient benzene showed an increase in 1996. The RFG compliance data (Table 2-7) also show that the fuel benzene content increased significantly from 1995 to 1996.

Many of the figures discussed above illustrate that benzene concentrations and/or weight fractions and ratios showed a step change between 1994 and 1995 with 1993 (if available) and 1994 values higher than 1995 and 1996 values. The t-tests and nonparametric tests confirmed the significance of the differences.

At the following sites that introduced RFG, the trend analysis results were mixed, with no clear downward trend in RFG indicators. Specific findings are as follows:

- Chicopee, MA (Springfield). Benzene concentrations and weight fractions, and xylene and trimethylbenzene fractions decreased between 1994 and 1995 (**Figure 4-10**). However, the benzene/toluene ratio increased. Also, the benzene weight percent returned to 1994 levels in 1996. The RFG compliance data (Table 2-7) show that additional declines in fuel aromatic content occurred between 1995 and 1996, but benzene levels were unchanged. Changes in non-mobile sources of toluene (for example) may be affecting ambient concentrations. Previous work by NESCAUM (1995) indicated that several hydrocarbons were being emitted by unknown local sources.

- East Hartford, CT. The benzene/toluene and benzene/acetylene ratios showed slight increases between 1994 and 1995. Fuel change data were not available for this city for this time period. Unfortunately, ambient data were not available for 1996. This analysis may be affected by the possibility of significant non-mobile sources of several hydrocarbons near this site (e.g., NESCAUM, 1995).

- Chicago, IL. Benzene concentrations showed a statistically significant decline between 1994 and 1995 with the introduction of RFG, however, the benzene/toluene and benzene/acetylene ratios showed increases during the same time period (**Figure 4-11**). It may be that the site and analytical changes between 1994 and 1995 offset the fuel changes in this time period. From 1995 to 1996, significant declines were observed in ambient benzene levels and the related ratios. In contrast, the RFG compliance data (Table 2-7) show that the fuel benzene content increased from 1995 to 1996.

4.2.3 Results for Sites That Did Not Implement RFG

At Baton Rouge, the site in this study without RFG implementation, benzene weight percent and the ratios showed statistically significant differences between 1994 and 1995 (**Figure 4-12**). However, there was a statistically significant increase in the benzene weight percent from 1995 to 1996 that was of a similar magnitude to the 1994 to 1995 decrease.

Milwaukee, WI

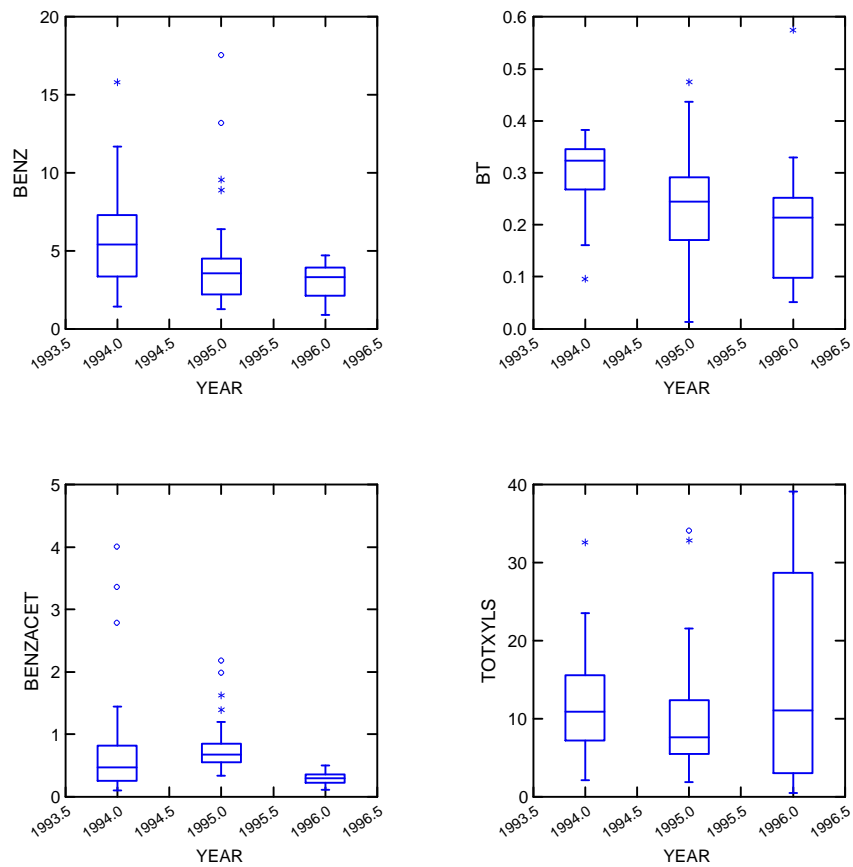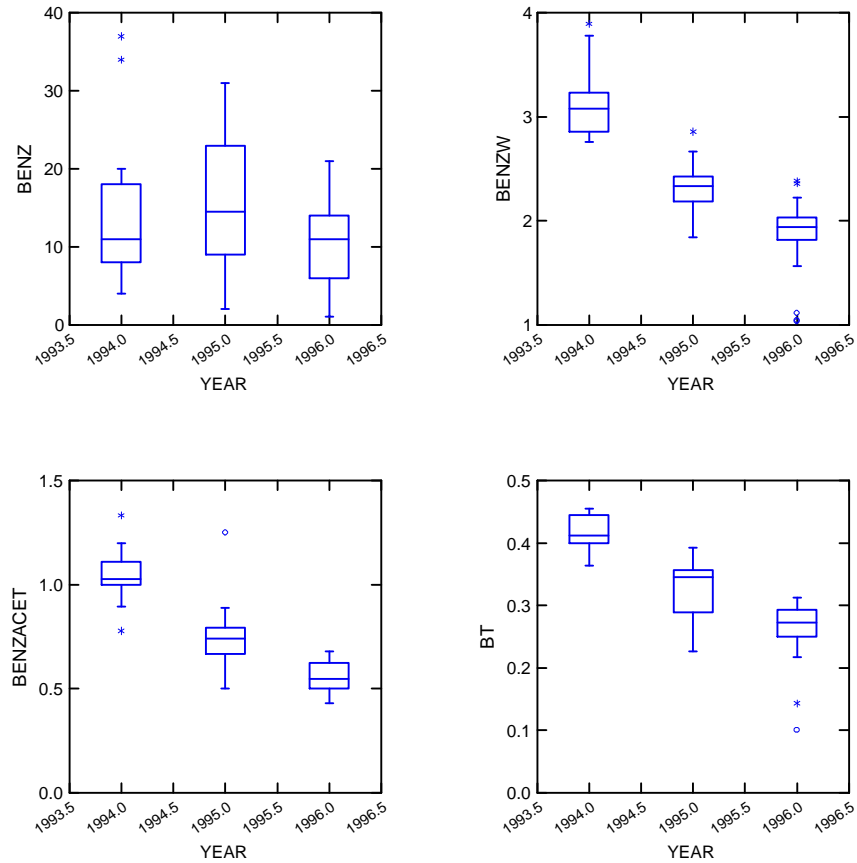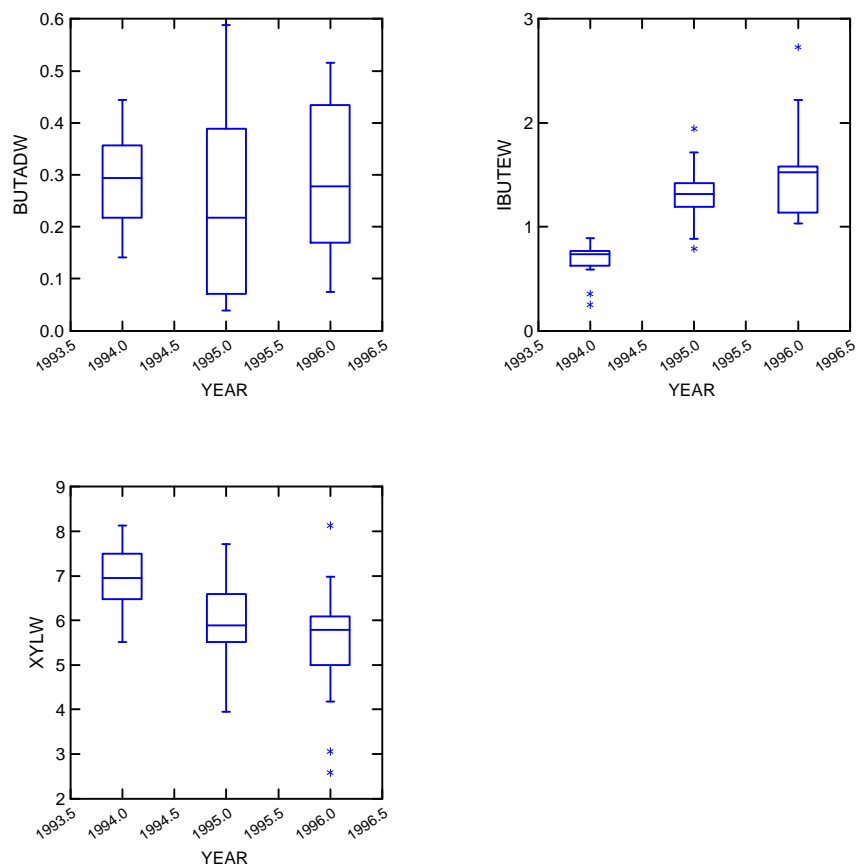

Figure 4-9.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent
(BENZW), (c) benzene to acetylene concentration ratio (BENZACET), and (d) n-
butane weight percent (NC4W) at Milwaukee, WI for the summers of 1994-1996.
Data are June-August, 0500-0800 CST, valid data only. Statistical tests indicated
significant differences between 1994 and 1995 and 1995 and 1996 benzene
concentrations.

Chicopee, MA
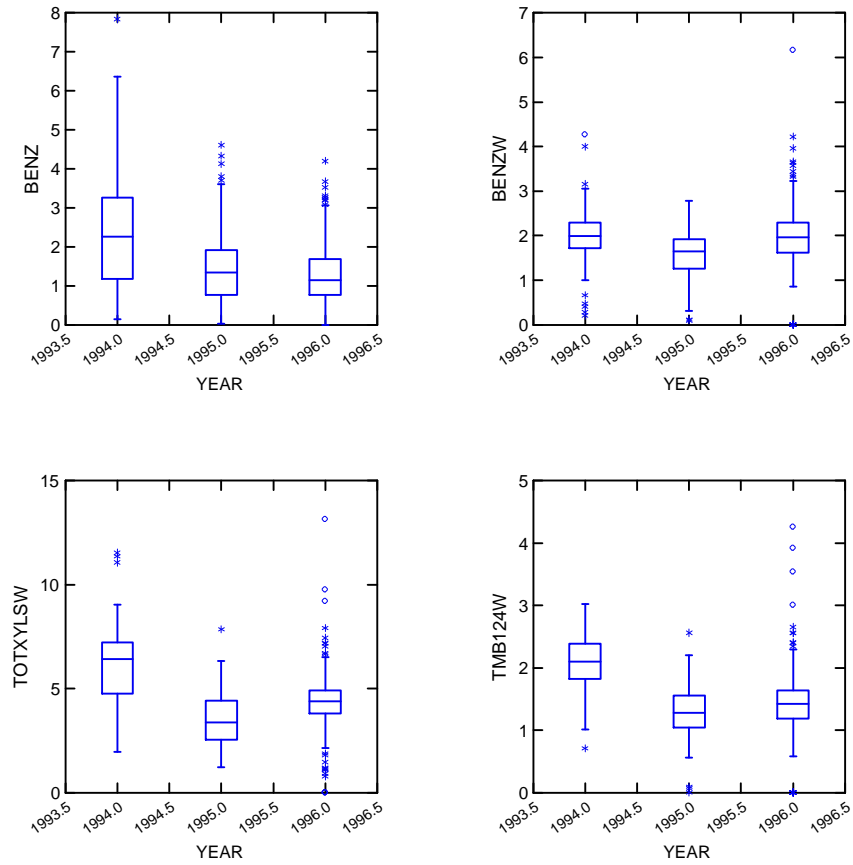
Figure 4-10.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent
(BENZW), (c) total xylenes weight percent (TOTXYLSW), and
(4) 1,2,4-trimethylbenzene weight percent (TMB124W) at Chicopee, MA for the
summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only.
The statistical tests showed that the 1994 and 1995 means were statistically
different for the species shown.

Chicago, IL

Figure 4-11.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent
(BENZW), (c) benzene to toluene concentration ratio (BT), and (d) benzene to
acetylene concentration ratio (BENZACET) at Chicago, IL for the summers of
1994-1996. Data are June-August, 0500-0800 CST, valid data only. Note that
the declines in benzene concentration and the benzene to toluene ratio occurred
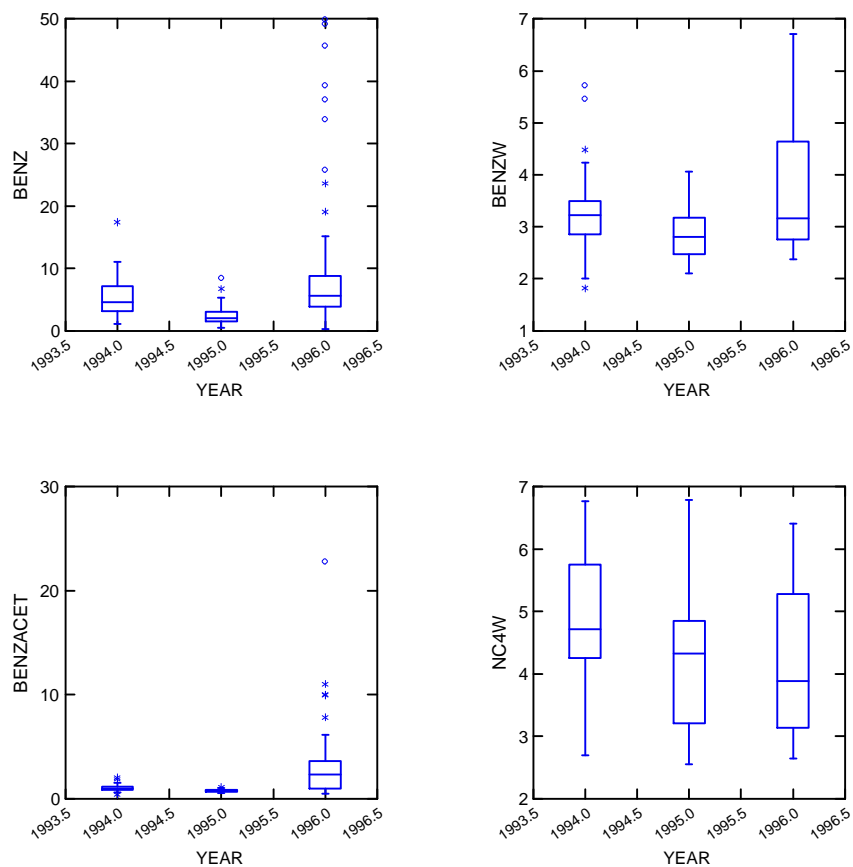between 1995 and 1996.

Figure 4-12.   Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent
(BENZW), (c) benzene to acetylene concentration ratio (BENZACET), and
(d) benzene to toluene concentration ratio (BT) at Baton Rouge, LA for the
summers of 1993-1996. Data are June-August, 0500-0800 EST, valid data only.

Benzene concentrations showed little change. Also, while benzene weight percents showed a decline between 1994 and 1995, the benzene/acetylene and benzene/toluene ratios increased. This is not what was observed at the other sites discussed above where a consistent picture of benzene declines was typically exhibited (e.g., in most cases a decline in benzene weight percent or concentration accompanied a decrease in the benzene ratios). Unfortunately, no fuel composition information was available for this city. Also, changes in local non-motor vehicle emissions may influence these results.

At Harrington Beach WI, the box plots showed mixed results with benzene concentrations declining between 1994 and 1995, but the benzene weight percent showed little change (**Figure 4-13**). Conversely, the benzene concentration showed a decrease in 1995 and 1996, but the weight percent in 1996 was significantly higher than in 1995. These mixed results may be due to the influence of transported Milwaukee emissions (where RFG was implemented) on the measurements at Harrington Beach.

4.3 RESULTS FOR 24-HR AVERAGE DATA

The analysis in Section 4.2 focused on the morning concentration data, since it was determined that this is the period where the affect of RFG is expected to be most apparent. In this section, we investigate whether the trend is also apparent in the 24-hour average data. It was hypothesized that there might be a significant weekday versus weekend difference in these data. However, exploratory plots of the data segregated by weekday versus weekend at Lynn, MA and Bronx, NY showed little difference in the results. Therefore, all data were used to increase the total sample population. **Figures 4-14 and 4-15** show box plots of the benzene concentration and weight percent and related ratios for 1993 through 1996 for the summer period (Figure 4-14) and the entire year (Figure 4-15). Both plots show statistically significant reductions for benzene from 1994 to 1995, consistent with the results obtained for the morning data. This shows that the RFG trend was also apparent in the 24-hour averaged data, both during the summer and throughout the entire calendar year.

4.4 INVESTIGATION OF ISOPRENE TRENDS

The analysis of the indicators selected to detect RFG has suggested that RFG indeed influenced ambient concentrations. If it could be shown that a compound not expected to be reduced with RFG did not show a trend over the RFG implementation period, this would provide further evidence of the influence of RFG. Fortunately, isoprene, a hydrocarbon of biogenic origin, is one of the PAMS target species and thus is available for this analysis. Isoprene emissions are a function of temperature and sunlight. Thus, isoprene concentrations and weight percent of total NMHC typically reach a maximum during midday or afternoon hours, and should vary day to day depending on the meteorological conditions. Peak isoprene concentrations have been found to vary widely from site to site depending upon the site's proximity to vegetation and other factors such as temperature, mixing height, and wind direction. For this analysis, the isoprene concentration and weight percent data at each site were investigated over the same time period as the RFG indicators.
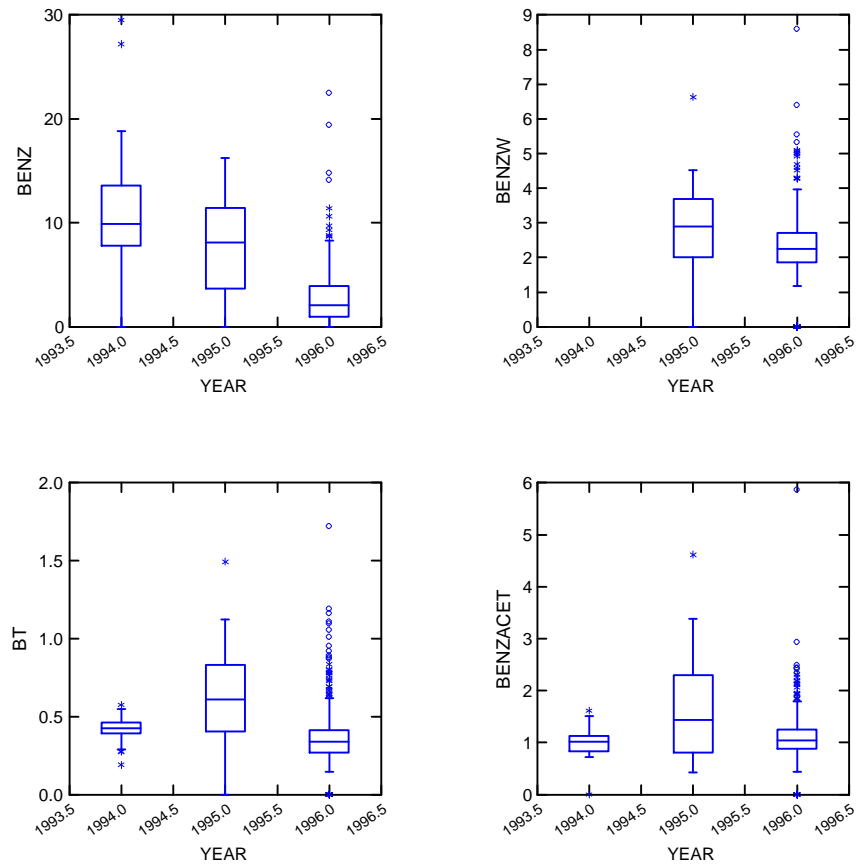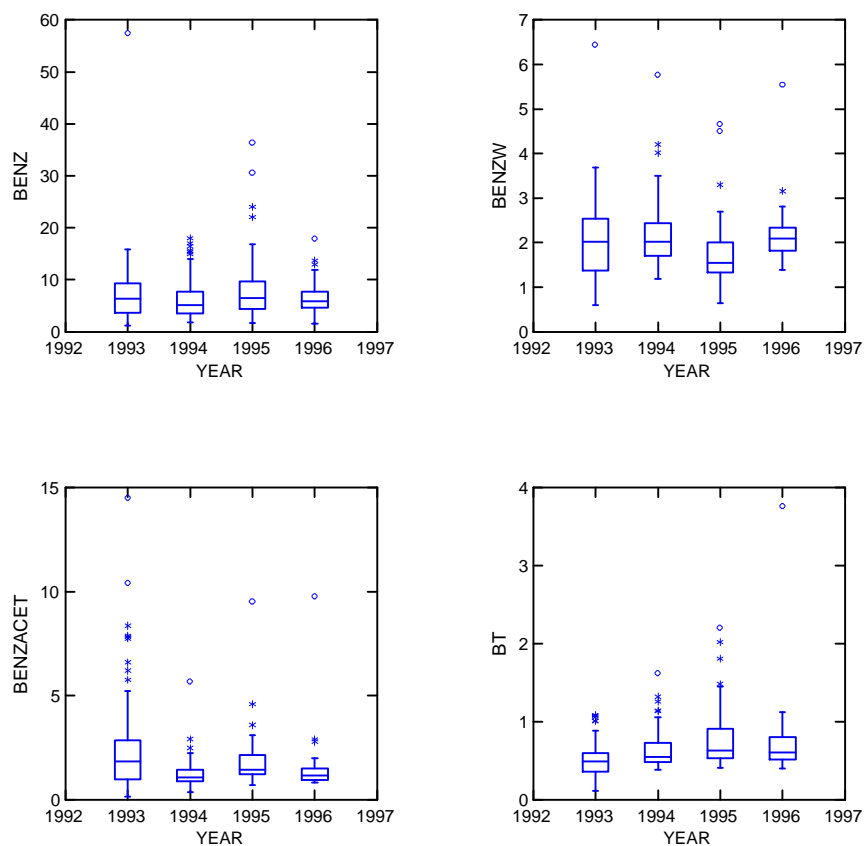
Harrington Beach, WI
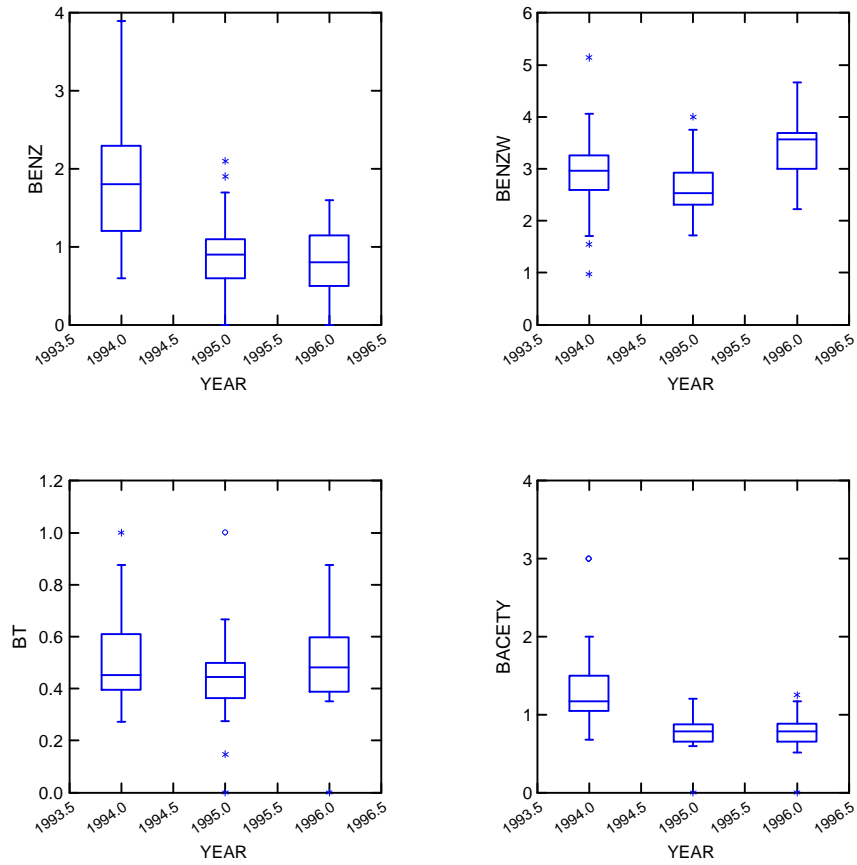
Figure 4-13. Box plots of (a) benzene concentration (BENZ), (b) benzene weight percent (BENZW), (c) benzene to toluene concentration ratio (BT), and (4) benzene to acetylene concentration ratio (BACETY) at Harrington Beach, WI for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only.

Lynn 24 hr (summer)
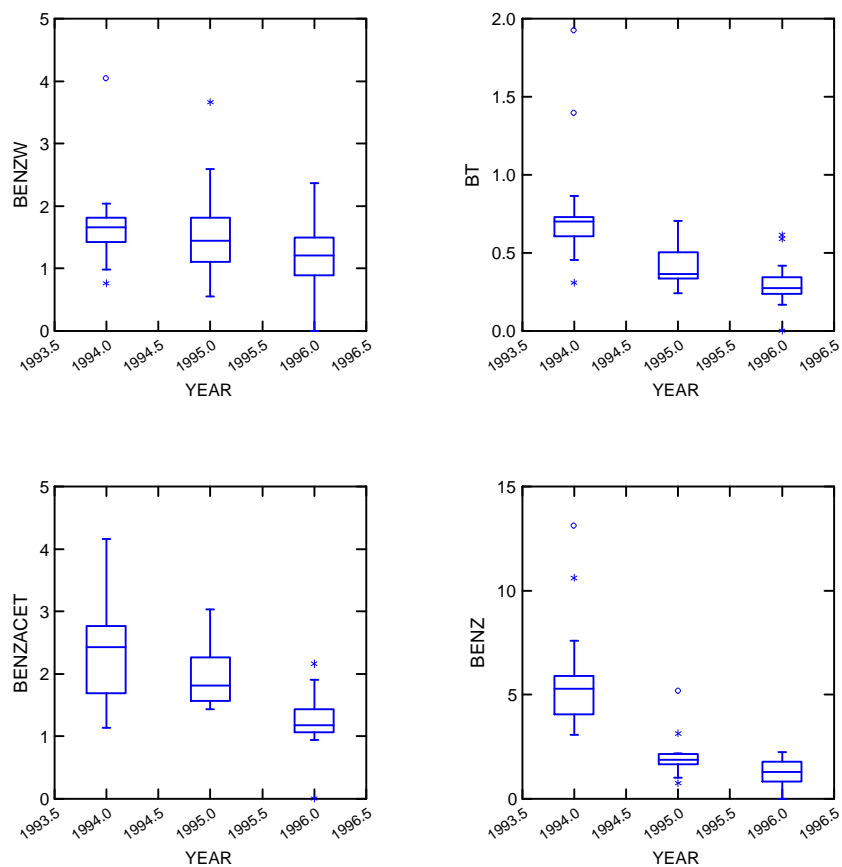
Figure 4-14.   Box plots of 24-hr average (a) benzene weight percent (BENZW), (b) benzene to toluene concentration ratio (BT), (c) benzene to acetylene concentration ratio (BENZACET), (d) benzene concentration (BENZ) at Lynn, MA for the summers of 1993-1996. Data for 1993 only included January through June and were omitted from these plots.

Lynn, 24-hr data (all year)
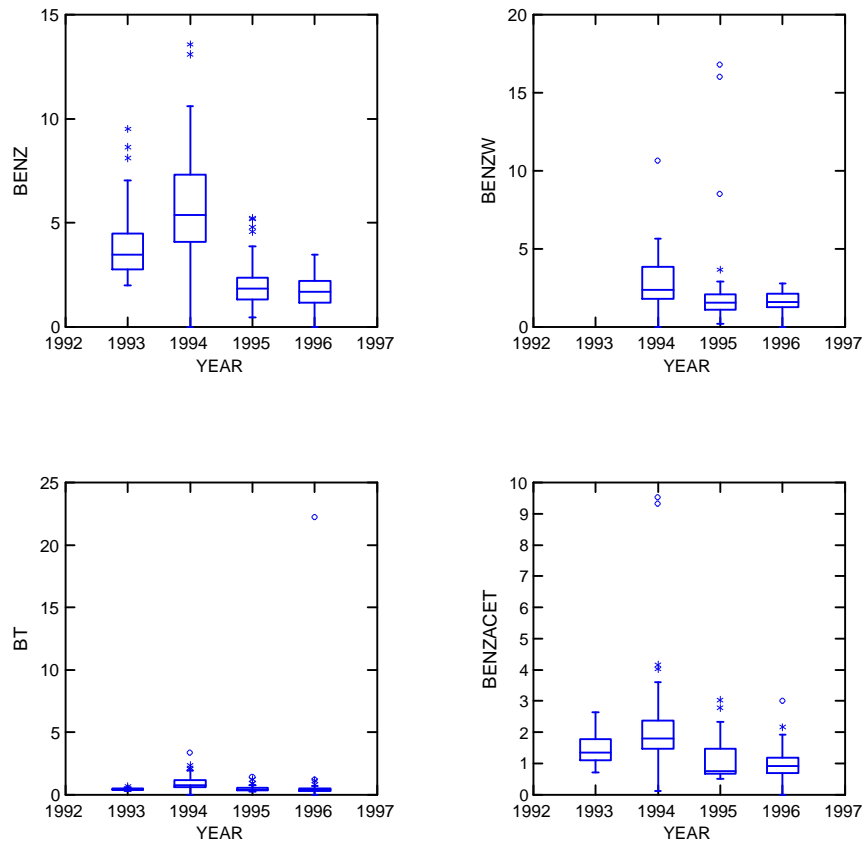
Figure 4-15. Box plot of 24-hr average (a) benzene concentration (BENZ), (b) benzene weight percent (BENZW), (c) benzene to toluene concentration ratio (BT), and (d) benzene to acetylene concentration ratio (BENZACET) at Lynn, MA for 1993-1996. Data for 1993 only included January through June.

4-23

Figures 4-16 through 4-18 show representative results from this analysis. These figures show that there is no discernible trend in the isoprene concentrations over the period that RFG was implemented. (Results in Tables 4-2 and 4-3 show mostly no change in isoprene over the study years.) Combined with the conclusion that the RFG indicators, such as the weight percent of benzene, did indeed decline over this same period, this analysis suggests that the trends in the RFG indicators were real and not due to other factors such as annual variations caused by meteorology.
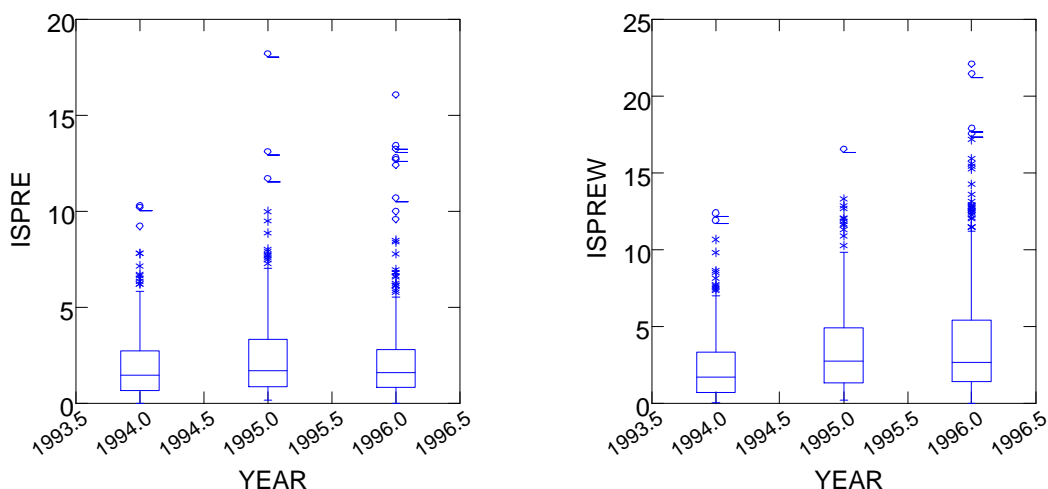
Figure 4-16. Box plots of isoprene concentration and weight percent at Chicopee, MA for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only. The statistical tests showed that the annual means were not statistically different.

McMillan Reservoir, DC
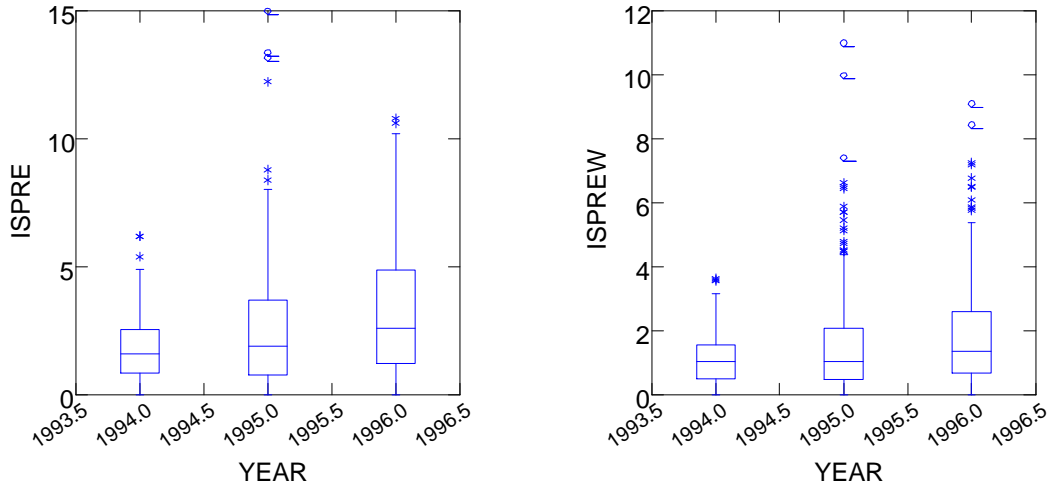
Figure 4-17. Box plots of isoprene concentration and weight percent at McMillan Reservoir (Washington, D.C.) for the summers of 1994-1996. Data are June-August, 0500-0800 EST, valid data only. The statistical tests indicated that the annual data were not different.

Baton Rouge, LA
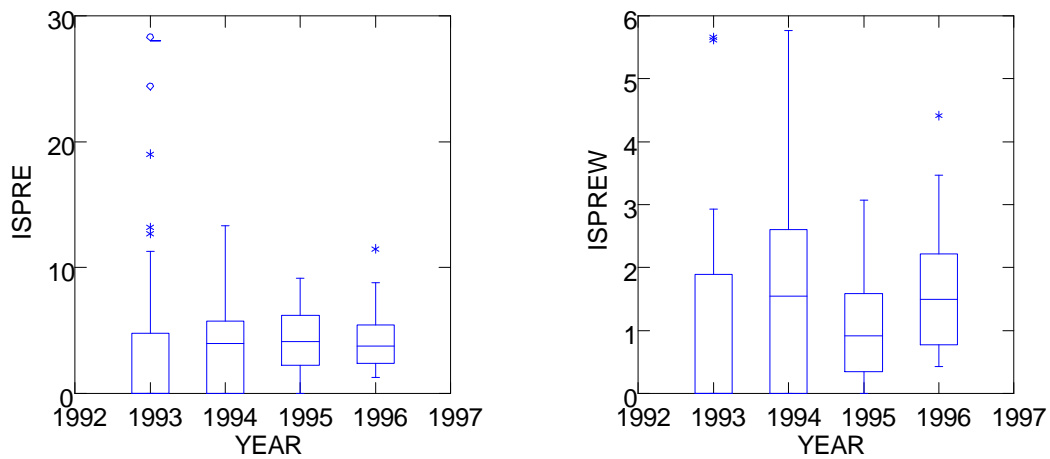
Figure 4-18. Box plots of isoprene concentration and weight percent at Baton Rouge, LA for the summers of 1993-1996. Data are June-August, 0500-0800 EST, valid data only.

4.5 METEOROLOGICAL INFLUENCES AND ADJUSTMENTS

Considerable work has been done to account for meteorological influences on ozone concentrations and adjust the data accordingly (e.g., Chinkin et al., 1996a; Rao et al., 1995; Cox

and Chu, 1993). However, less research has focused on meteorological influences and adjustments to hydrocarbon data. Just as for ozone, any meteorological variable that influences the development and height of the mixing layer or the potential for transport is likely to influence the concentrations of hydrocarbons. Furthermore, the temperature is likely to directly influence emission rates due to factors such as increased energy demand and increased evaporative emissions during hotter periods. For hydrocarbons, it is likely that the following meteorological parameters may influence (directly or indirectly as they correlate with the height of the mixing layer) concentrations: (1) temperature, (2) wind direction, and (3) wind speed. When these meteorological parameters vary from year to year, it is necessary to adjust for their effect on concentrations.

One of the simplest adjustment factors for the effect of meteorology on the number of ozone exceedances is the number of days above 90°F. To apply this technique to PAMS VOC trends, the average annual weight fraction or concentration of the VOC can be divided by the number of days above 90°F, and the trend re-analyzed using the above mentioned methods.

A more sophisticated technique would be to use time-series analysis. This analysis would be based on the trend in the daily index and not the averages for each year. This is more accurate than using annual averages because the influence of meteorology is daily. Cox and Chu (1993) have developed a regression model for predicting the daily ozone concentration based on the values of several meteorological variables and a term that accounts for the annual trend. The following similar regression model is proposed for the PAMS trends:

$$VOC = \mathrm{b}_o + \mathrm{b}_1 Temperature + \mathrm{b}_2 Wind\ Speed + \mathrm{b}_3 X_i$$

where X_i is equal to 1 for years after RFG was implemented, and 0 for years before RFG was implemented. The X_i term represents the meteorologically adjusted changes in the VOC index that occurred when RFG was implemented. VOC is the appropriate VOC index for a particular day. The meteorological variables correspond to the morning hydrocarbon values of those parameters. It is worth noting that most PAMS stations collect all the necessary meteorological data. This is usually not the case with ozone stations, where meteorological data sometimes needs to be taken from nearby National Weather Service (NWS) stations. Temperature is likely to be the meteorological variable that best predicts the VOC weight fraction or concentration, but wind speed may also be important, particularly at the Type III (maximum ozone) and Type IV (far downwind) sites. Wind direction may also play a role. However, it is difficult to incorporate wind direction into a regression model because it is not a linear variable (i.e., it is bounded from 0 to 360 degrees depending on the direction).

For the regression model above, the benzene weight percent in 1994 and 1995 will be used. The results for the regression model for Lynn, MA and McMillan Reservoir, D.C. are shown in **Tables 4-5 and 4-6**. For the Lynn, MA site, both temperature and wind speed were

Table 4-5.   Results of the meteorological-adjustment regression model for Lynn, MA. The model tests whether the benzene weight percent declined from 1994 to 1995.

Model Variable	Regression Coefficient	t-statistic	P-value
Constant	4.3	9.7	0.000
Temperature	-0.083	-4.4	0.000
Wind Speed	0.105	3.4	0.011
Year	-1.44	-9.3	0.000

Table 4-6.   Results of the meteorological-adjustment regression model for McMillan Reservoir, D.C. The model tests whether the benzene weight percent declined from 1994 to 1995.

Model Variable	Regression Coefficient	t-statistic	P-value
Constant	3.8	7.3	0.000
Temperature	-0.017	-0.8	0.43
Wind Speed	0.039	0.7	0.48
Year	-1.24	-8.7	0.000

statistically significant, suggesting that meteorology influenced the daily fluctuations in the benzene weight percent. However, the dummy variable for year was also highly significant and negative. This shows that the benzene weight percent declined from 1994 to 1995. The value of the year variable was -1.4. This suggests that the meteorologically-adjusted reduction was 1.4 weight percent benzene from 1994 to 1995, while the observed reduction was only 1.0 weight percent benzene. For the McMillan Reservoir, D.C. site, the meteorological variables were not significant. However, the dummy variable for year was statistically significant, which shows that the benzene weight percent declined from 1994 to 1995. This is highly consistent with the observed reduction of 1.3 weight percent benzene, which is expected since the meteorological variables had no effect.

These calculations show that the statistical results for two of the sites still hold after adjustment for meteorology. Furthermore, in the Lynn, MA case, the meteorological variables were statistically significant which shows that fluctuations in meteorology explained part of the fluctuations in benzene weight percents. If the observed differences in benzene weight percent had been smaller, this meteorological-adjustment technique may have proven to be more important to the analysis. Limited resources prevented doing this calculation for all of the sites, but this may be an area for future work.

4.6 OTHER ANALYSIS METHODS

The implementation of PAMS data collection in the major ozone nonattainment areas of the United States provides a valuable data set for assessing the impacts of RFG as well as tracking air quality trends. However, the number of PAMS sites is limited and of only limited duration at this time. Thus, it is important to consider the benefits of using auxiliary data sets to complement the available PAMS data. Example auxiliary data that could be quite beneficial include data collected in successive years in the same location such as VOC and NO_x measurements made in tunnels in California in both the Los Angeles (Gertler and Sagebiel, 1997) and San Francisco areas (Kirchstetter et al., 1997). In some ways these tunnel studies offer the potential to more precisely assess emissions changes due to RFG than PAMS sites since they are measurements of predominately motor vehicle exhaust emissions, prior to and subsequent to RFG implementation. For example, Kirchstetter et al. (1997) found that benzene and nonmethane organic compounds (NMOC) concentrations decreased as measured during 1994, 1995, and 1996 in the Caldecott tunnel. The authors note that their results should be examined to assess the possible effects of fleet turnover.

Another independent means of assessing RFG benefits is to compute expected air quality changes from estimated emission changes using either simple mathematical models such as EKMA, or more sophisticated models such as the Urban Airshed Model (UAM). While the EKMA is simple to use, the results of the EKMA may not truly reflect the air quality changes expected because emission changes resulting from RFG are likely to be spatially and temporally distributed in a nonhomogeneous fashion (a violation of key EKMA modeling assumptions). While more sophisticated three-dimensional modeling such as UAM can account for spatially and temporally allocated emissions, it is often difficult to assess how much of the changes in predicted air quality can be attributed to emission changes as opposed to model uncertainties resulting from wind field and chemistry artifacts of the modeling process. Another potential problem with emissions-based modeling is the designation of the appropriate speciated source profiles for pre- and post-RFG motor vehicle exhaust. Recent work at STI for the Coordinating Research Council has found some problems regarding the contents of various VOCs including MTBE and benzene in the exhaust profiles.

4.7 CONCLUSIONS

This section provides the results of the RFG trend analysis for the selected metropolitan areas. Several indicators expected to change with the introduction of RFG are listed in Table 4-1. Whenever PAMS data were available, these indicators were calculated using the morning concentrations. The results of this analysis can be summarized as follows:

- The following sites showed reductions among the indicator variables that were consistent with the implementation of RFG in the area: McMillan Reservoir, D.C.; East Providence, RI; Philadelphia, PA; Bronx, NY; and Los Angeles, CA. We conclude that the ambient concentration changes expected with RFG were realized at these sites.

- For Lynn, MA and Milwaukee, WI, ambient benzene concentrations decreased between 1994 and 1995 (as did fuel benzene content) but increased in 1996. The RFG compliance data showed that the fuel benzene content increased significantly from 1995 to 1996 in

these areas. Thus, while the benzene trends were mixed, the ambient data and fuel composition data agreed. It is reasonable to infer that year to year variations in RFG fuel properties contribute to year to year changes in ambient levels of benzene and other RFG indicators.

- The following sites showed mixed results for the trend analysis, even though RFG was implemented in the area: Chicopee, MA; East Hartford, CT; and Chicago, IL. At these sites, benzene concentration and weight percent data changes differed from the benzene-related ratio changes. In addition, fuel composition data did not appear to be consistent with the ambient changes. We are uncertain that the ambient concentration changes expected with RFG were realized. Other sources of RFG indicator species, analytical or sampling changes, or meteorological variability may compound this analysis.

- The following sites showed mixed results for the trend analysis: Baton Rouge, LA and Harrington Beach, WI. At these sites, benzene concentration or weight percent changes from 1994 to 1995 were offset by changes from 1995 to 1996. In addition, benzene ratio changes were in the opposite direction (or did not show a change). These areas did not implement RFG. Fuel content information was not available for these sites to corroborate the ambient changes. We are uncertain of the causes of ambient concentration changes during the RFG implementation period at these sites.

Additional analyses showed that the results were consistent when using 24-hour average data at Lynn, MA, both during the summer and over the entire year.

There are other factors besides RFG that can cause changes or fluctuations in the ambient concentration. These include other hydrocarbon sources and meteorological parameters such as temperature and wind speed. Additional analyses were performed to investigate the influence of these parameters:

- A meteorological adjustment model was developed for the benzene weight percent based on temperature and wind speed. For Lynn, MA the model showed that the temperature and wind speed explained some of the variability in benzene weight percent, but the benzene weight percent trend was still apparent. For McMillan Reservoir, D.C., the meteorological variables were not significant. In both cases, there was still significant trend in benzene indicators, even when meteorological influences were considered.

- The trend in the biogenic compound isoprene was investigated. Isoprene concentrations are known to be influenced by meteorology, but should not decline with the implementation of RFG. This analysis showed that while there was a decline in ambient benzene concentrations at many sites, there was no trend (or a slight positive trend) in isoprene concentrations over the RFG implementation period. Combined with the conclusion that the RFG indicators, such as the weight percent of benzene, did indeed decline over this same period, this analysis suggests that the trends in the RFG indicators were real and not due to other factors such as annual variations caused by meteorology.

5. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE ANALYSES

The two primary objectives of this project were to determine (1) if the PAMS measurements indicate real reductions in ambient concentrations of hydrocarbons, and (2) to what extent these reductions in ambient concentrations can be linked to RFG implementation in early 1995. This project entailed considerable data gathering: RFG implementation schedules for the various areas, fuel composition information for 1993-1996, identification of potential indicators to show RFG effects, and PAMS site and sampling details. Thirteen PAMS data sets, encompassing areas with and without RFG implementation, were obtained, validated, and investigated. Fuel surveys and RFG implementation plans showed that most of the areas with PAMS sites had implemented RFG during early 1995 including Chicago, IL; Boston (Lynn) and Springfield (Chicopee), MA; Milwaukee, WI; Philadelphia, PA; Bronx, NY; Los Angeles, CA; E. Providence, RI; Washington, D.C. (McMillan Reservoir); and E. Hartford, CT. Baton Rouge, LA and Harrington Beach, WI were two PAMS sites that were in areas that did not implement RFG. Section 5.1 provides the conclusions drawn from the investigation of the PAMS data. Section 5.2 provides recommendations for further work.

5.1 CONCLUSIONS

One of the questions to be addressed in this project was whether or not the PAMS data were sufficient, both in quantity and quality, for a trend analysis. Only two sites had a significant amount of data available from 1993; therefore, in most cases, the pre-RFG data were limited to 1994. While we noted various problems with the data quality in Section 3 of this report, in general, the data were sufficient for further investigation after careful and thorough quality control/validation. In addition, by selecting indicators such as the ratio of benzene to toluene, data deficiencies, such as a lack of valid NMHC values, could be overcome.

Fuel Changes

The first step in the trend analysis was to identify what changes were made to gasoline to meet the RFG requirements. This proved to be more difficult than expected, due to the paucity of fuel composition data prior to the implementation of RFG. As part of the RFG implementation, petroleum companies were required to submit fuel composition data to the EPA. Thus, for the post-RFG period, there were generally adequate fuel composition data for most areas that were investigated. The following conclusions were made from the analysis of the fuel data:

- Fuel survey data generally showed that RFG implementation resulted in a reduction in RVP, fuel aromatic and olefin content, and benzene. Increases in fuel saturates and MTBE were also observed.

- Benzene, which was targeted for specific reductions in the RFG implementation, showed changes of up to 1.0 weight percent of the fuel composition from 1994 to 1995.

- Fuel data were not available for several cities with a PAMS Type II (emission source) monitor that were investigated in this report, including E. Providence, RI; Baton Rouge, LA; and Milwaukee, WI.

Ambient Changes

The next step in the analysis was to confirm that the changes observed in fuel composition actually resulted in changes in the ambient concentration. This was done through an examination of PAMS data at sites in areas which implemented RFG and some that did not implement RFG.

Several indicators were chosen to evaluate the expected trends due to RFG. The most important were the benzene concentration and weight percent (i.e., the benzene concentration divided by the NMHC). Since NMHC was not always reported at PAMS sites early in the program, the ratio of the benzene-to-toluene concentrations and the benzene-to-acetylene concentrations were also calculated. Toluene is not expected to change as much as benzene and is one of the compounds that is most reliably measured at PAMS sites; and acetylene is not expected to change due to RFG (although other emission sources can contribute to toluene or acetylene). Since RFG also requires a reduction in total aromatics, the trend in several aromatic species were tested including total xylenes and trimethylbenzenes. MTBE is not measured at PAMS sites. The olefin i-butene, which is a thermal decomposition product of MTBE, is not on the PAMS target species list, but was measured at Los Angeles and the trend at this site was evaluated.

It is expected that meteorological factors cause day-to-day fluctuations in hydrocarbon concentrations and can also cause annual variations independent of the fuel composition. Therefore, a meteorological adjustment model was also developed to test if the meteorologically-independent concentration of benzene was reduced. The trend for isoprene, a biogenic compound, was evaluated because this is a compound with a concentration that fluctuates with meteorology but is not expected to decline with RFG.

The trends in these indicators were evaluated using standard statistical tests that are described in the report. The following conclusions were made about each of the metropolitan areas:

- The following sites showed reductions among the indicator variables that were consistent with the implementation of RFG in the area: McMillan Reservoir, D.C.; East Providence, RI; Philadelphia, PA; Bronx, NY; and Los Angeles, CA. We conclude that the ambient concentration changes expected with RFG were realized at these sites.

- For Lynn, MA and Milwaukee, WI, ambient benzene concentrations decreased between 1994 and 1995 (as did fuel benzene content) but increased in 1996. The RFG compliance data showed that the fuel benzene content increased significantly from 1995 to 1996 in these areas. Thus, while the benzene trends were mixed, the ambient data and fuel composition data were consistent. It is reasonable to infer that year to year variations in

RFG fuel properties contribute to year to year changes in ambient levels of benzene and other RFG indicators.

- The following sites showed mixed results for the trend analysis, even though RFG was implemented in the area: Chicopee, MA; East Hartford, CT; and Chicago, IL. At these sites, benzene concentration and weight percent data changes differed from the benzene-related ratio changes. In addition, fuel composition data did not appear to be consistent with the ambient changes. We are uncertain that the ambient concentration changes expected with RFG were realized. Other sources of RFG indicator species, analytical or sampling changes, or meteorological variability may compound this analysis.

- The following sites showed mixed results for the trend analysis: Baton Rouge, LA and Harrington Beach, WI. At these sites, benzene concentration or weight percent changes from 1994 to 1995 were offset by changes from 1995 to 1996. In addition, benzene ratio changes were in the opposite direction (or did not show a change). These areas did not implement RFG and fuel content information was not available for these sites to corroborate the ambient changes. We are uncertain of the causes of ambient concentration changes during the RFG implementation period at these sites.

The following conclusions can be made about the trends observed in each of the compounds examined:

- Significant decreases, from 18 to 42 percent, in ambient benzene median weight percents occurred at most sites between 1994 and 1995 (summer morning data). The benzene ratios (benzene to toluene and benzene to acetylene) also declined. At all the sites that showed a decrease in benzene in ambient air, there was a decrease in fuel benzene content recorded as well.

- Aromatic indicators, such as total xylenes and trimethylbenzenes, also showed significant declines at most of the sites.

- The ambient i-butene weight percents increased significantly between 1994 and 1995 in Los Angeles (the only site that reported an adequate number of samples with this species above detection) as might be expected since i-butene is a thermal decomposition product of MTBE.

- In contrast to the trends in ambient benzene observed at the sites, ambient summer morning isoprene (a biogenic hydrocarbon) concentrations and weight percents did not show a significant change from 1994 to 1995. This is further evidence that the changes in the ambient benzene were not due to meteorological effects.

When one considers the above findings, a strong case emerges that supports the observation that a reduction of ambient benzene was related to the reduction of benzene in the fuel, and that there may also be significant reductions of other species such as aromatics. However, the results were not consistent across all of the sites. These inconsistencies may be due

to inaccurate fuel data to document the expected changes with RFG implementation or with measurement problems at the PAMS sites.

5.2 RECOMMENDATIONS

5.2.1 Recommendations for the PAMS Program

Future analyses of these trends depends on high-quality PAMS data. Therefore, the following recommendations are made for the PAMS program.

- Future PAMS measurements should include routine measurements of the oxygenate used in the region's RFG and/or their decomposition products.

- Future PAMS measurements should include routine measurements of other species that were predicted to increase with RFG implementation such as i-butene and 1,3-butadiene. Changes may need to be made in the measurement systems in order to accurately measure these species.

- The reporting agencies should carefully validate their PAMS data prior to submittal to AIRS so that calibration data, species misidentifications, and data resulting from instrument problems are not submitted as ambient data.

- If erroneous data have been submitted, the reporting agencies should update their AIRS databases to **remove or correct** the data as soon as they have been identified (either by the reporting agency or by others).

- The reporting agencies need to consistently provide the total nonmethane hydrocarbon values. If the NMHC is provided along with the individual species, the unidentified portion of the NMHC is calculable. Reporting agencies should add these data to AIRS where they are missing.

- The reporting agencies should report other species in addition to the PAMS target list if those species are important at the site. Some species may be unique (or nearly so) to different sources.

- The begin times of samples need to be accurately reported in standard time. Reporting agencies should correct data already incorrectly reported in AIRS to reflect standard time.

5.2.2 Suggestions for Further Work

The analyses in this report build on the existing methodology for the analysis of trends in ambient hydrocarbon concentrations as the result of fuel changes. The following suggestions pertain to future analyses.

- Further investigate the effect of transport and background VOC on this analysis. Determine screening criteria using ratios of more-reactive to less-reactive hydrocarbons (this assesses the age of the air mass) and concentrations. Minimizing transport and background influences should increase the accuracy of the results shown in this report.

- Look at benzene and other indicator trends for other time periods during the day. However, it is likely that afternoon trends will be more influenced by changes in meteorology and chemical reactions.

- A careful comparison of meteorology and the ambient VOCs could be made for sites that changed analytical techniques, sampling interval, and/or location, in order to increase the available database for this type of analysis.

- Investigate other possible changes (such as RVP reductions unrelated to RFG) that could alter benzene concentrations.

- Further assess effects of meteorology on apparent ambient trends, particularly at non-RFG sites such as Baton Rouge. Consideration of sampling intervals will need to be made (i.e., 1-hr meteorological data versus 3-hr average hydrocarbon data).

- Investigate recently performed (i.e., in November 1997) work in California tunnels by U.C. Berkeley researchers in context with their measurements in 1994-1996. How do these results compare to the results in this report? Do the tunnel studies provide more information regarding the effects of fleet turnover or other processes that may influence ambient concentrations?

- At this time, it is unlikely that more complicated modeling, such as using the UAM, would provide more meaningful results than those obtained here. Uncertainties in source profiles for pre-and post-RFG motor vehicle exhaust and evaporative emissions need to be addressed before pursuing this type of modeling endeavor.

- In the 1994-1996 time frame, changes in fuel benzene have been relatively large and this study showed that the changes were relatively easy to see in ambient air using simple statistical techniques. However, in the future, more modest changes in fuel composition may be more difficult to detect in the ambient air. There are several methods in the current literature that could be applied to investigate harder-to-detect trends such as Larsen's "bootstrap" method (Larsen et al., 1990) and Rao and Zurbenko's filtering techniques (e.g., Rao and Zurbenko, 1994).

6. REFERENCES

Chinkin L.R., Reiss R., Eisinger D.S., Dye T.S., and Jones C.M. (1996) Ozone exceedance data analysis: representativeness of 1995. Phase I. Final report prepared for American Petroleum Institute, Washington, DC by Sonoma Technology, Inc., Santa Rosa, CA, STI-996031-1574-FR, August.

Cox, W.M. and Chu, S.H. (1993) Meteorologically adjusted ozone trends in urban areas: a probabilistic approach, *Atmos. Environ.*, **27B**, 425-434.

Gertler A. and Sagebiel J. (1997) 1987 to 1995 - how have mobile source emissions changed in the Van Nuys tunnel? Paper presented at the *7th CRC On-Road Vehicle Emissions Workshop, San Diego, CA, April 9-11.*

Hammond D. (1996) Ambient trends of benzene in California from 1990 through 1995. Paper presented at the *U.S. Environmental Protection Agency and Air & Waste Management Association International Symposium on Measurement of Toxic and Related Air Pollutants, Research Triangle Park, NC, May 7-9.*

Kirchstetter T.W., Singer B.C., Harley, R.A. Kendall G.R., and Traverse M. (1997) Impact of phase 2 reformulated gasoline on California vehicle emissions. Paper presented at the *7th CRC On-Road Vehicle Emissions Workshop, San Diego, CA, April 9-11.*

Larsen L.C., Bradley, R.A., and Honcoop G.L. (1990) A new method of characterizing the variability of air quality-related indicators. Paper presented at the *Air & Waste Management Association International Specialty Conference on Tropospheric Ozone and the Environment, Los Angeles, CA, March 21.*

Ligocki M. and Yarwood G. (1994) RFG Indicators. Memorandum from Systems Applications International to CRC/NERL feasibility study technical committee, March 23.

Main H.H. and Roberts P.T. (1994) Investigation of hydrocarbon data in the South Coast Air Basin. Draft final report prepared for Systems Applications International, San Rafael, CA and the Coordinating Research Council by Sonoma Technology, Inc., Santa Rosa, CA, STI-94260-1415-DFR2, June.

Main H.H., Roberts P.T., and Korc M.E. (1997) PAMS data analysis workshop: illustrating the use of PAMS data to support ozone control programs. Prepared for U.S. Environmental Protection Agency, Research Triangle Park, NC, presented at EPA Region V Headquarters, Chicago, IL by Sonoma Technology, Inc., Santa Rosa, CA, STI-997100-1708-WD5, March.

Main H.H., Roberts P.T., Prouty J.D., and Korc M.E. (1996) Software for display, quality control, and analysis of continuous VOC data. Report prepared for Electric Power Research Institute, Palo Alto, CA by Sonoma Technology, Inc., Santa Rosa, CA, STI-996142-1594, EPRI Research Project No. WO9108-01, June.

Main H.H., Roberts P.T., and Ligocki M.P. (1995) Assessing the usefulness of VOC data as indicators of change in fuel composition in the South Coast Air Basin (California). Paper No. 95-FA113C.02 presented at the *Air & Waste Management Association 88th Annual Meeting, San Antonio, TX, June 18-23*.

NESCAUM (1995) Preview of the 1994 ozone precursor concentrations in the northeastern U.S. 5/1/94 draft report prepared by the Ambient Monitoring and Assessment Committee of the Northeast States for Coordinated Air Use Management, Boston, MA.

Poore, M. (1997) Personal communication.

Rao, S.T. and Zurbenko, I.G. (1994) Detecting and tracking changes in ozone air quality. *J. Air & Waste Manage. Assoc.*, **44**, 1089-1092.

Rao, S.T., Zalewsky, E., and Zurbenko, I.G. (1995) Determining spatial and temporal variations in ozone air quality, *J. Air & Waste Manage. Assoc.*, **45**, 57.

Stoeckenius, R.T., Yarwood, G., Ligocki, M.P., Cohen, J.P., Shepard, S.B., Looker, R.E., Fujita, E.M., Main, H.H., Roberts, P.T. (1995) Feasibility study for a 1995-1996 Southern California air quality monitoring program. Final report prepared for Coordinating Research Council, Atlanta, GA, by Systems Applications International, San Rafael, SYSAPP94-94/065d.

Stoeckenius, T.E., Yarwood, G., and Shepard, S.B. (1996) Development of PAMS data analysis techniques with application to Baton Rouge, Louisiana. Report prepared by ENVIRON International Corporation, Novato, CA, September.

U.S. Environmental Protection Agency (1984) Quality assurance handbook for air pollution measurement systems, volume ii: ambient air specific methods (interim edition), EPA/600/R-94/0386, April.

U.S. Environmental Protection Agency (1994) Photochemical assessment monitoring stations implementation manual. Report prepared by Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-454/B-93-051, March.

U.S. Environmental Protection Agency (1996) Air quality and emissions trends report, 1995. Report prepared by Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Walpole, R.E. and Myers, R.H. (1989) Probability and statistics for engineers and scientists, 4th Edition, McMillan Publishing Company.

Zielinska, B., Shire J., Harshfield G., Sagebiel J., Pasek R. (1997) VOC concentrations in the South Coast Air Basin (CA) following introduction of reformulated gasoline. Paper presented at the *Air & Waste Management Association Measurement of Toxic and Related Air Pollutants Symposium Conference*, Raleigh Durham, NC, April 29-May 1.

APPENDIX A

SUMMARY OF PAMS HYDROCARBON DATA VALIDATION RESULTS

Table A-1. Summary of suspect and invalid samples in the 1993 June-August Lynn, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Lynn, MA	250092006	930616	0300	I	uidvoc, NMHC	1	negative value
Lynn, MA	250092006	930618	0000	I	uidvoc, NMHC	1	negative value
Lynn, MA	250092006	930615	1000, 1100	I	sample	2	Low concentration
Lynn, MA	250092006	930604	1500	I	sample	1	Low concentration
Lynn, MA	250092006	930618	0600	I	sample	1	Low concentration

[a] Problems with identification of 2- and 3-methylpentane. Only two weeks of data.

Table A-2. Summary of suspect and invalid samples in the 1994 June-August Lynn, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Lynn, MA	250092006	940702	0400	S	ethyl	1	High concentration
Lynn, MA	250092006	940629	0100	S	sample	1	Calibration or peak reference data
Lynn, MA	250092006	940606	1600-2300	S	sample	8	Misidentification problems
Lynn, MA	250092006	940607	0000-2300	S	sample	23	Misidentification problems
Lynn, MA	250092006	940701	1400	S	sample	1	High nonane, other C8+ species
Lynn, MA	250092006	940702	1100	S	sample	1	High nonane, other C8+ species
Lynn, MA	250092006	940628	1600	S	sample	1	High nonane, other C8+ species
Lynn, MA	250092006	940625	2200	S	sample	1	Misidentification problems
Lynn, MA	250092006	940608	0000-1200	S	sample	12	Misidentification problems
Lynn, MA	250092006	940706	1900	S	sample	1	Misidentification problems
Lynn, MA	250092006	940728	1400	S	nnon	1	High concentration
Lynn, MA	250092006	940719	1400, 1500	S	nnon, ndec	2	High concentration
Lynn, MA	250092006	940708	1300	S	nnon	1	High concentration
Lynn, MA	250092006	940708	0500	S	npnta, ispna	1	Zero concentration
Lynn, MA	250092006	940705	1400, 1500	S	nnon	2	High concentration
Lynn, MA	250092006	940824	1000-2300	S	samples	13	No C2-C6
Lynn, MA	250092006	940825	0000-2200	S	samples	22	No C2-C6
Lynn, MA	250092006	940826	0000-2300	S	samples	23	No C2-C6
Lynn, MA	250092006	940827	0000-2300	S	samples	22	No C2-C6
Lynn, MA	250092006	940828	0000-1800	S	samples	18	No C2-C6
Lynn, MA	250092006	940803	1400	S	samples	1	No C2-C6
Lynn, MA	250092006	940811	1400	S	samples	1	Calibration or peak reference data
Lynn, MA	250092006	940715	1200, 1300	S	nnon, ndec	2	High concentration
Lynn, MA	250092006	940727	0900	S	sample	1	Missing key species

[a] Noted several samples with calibration carryover (i.e., high concentrations of n-nonane and n-decane) in July, several 2-methylhexane spikes, missing NMHC values during several periods.

Table A-3. Summary of suspect and invalid samples in the 1995 June-August Lynn, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments
Lynn, MA	250092006	950602	1300,1700, 1800	S	sample	3	High concentrations of 124tmb, unidentified
Lynn, MA	250092006	950604	0100-0300, 0600, 0900-1300, 2200	S	sample	10	High concentrations of 124tmb, unidentified
Lynn, MA	250092006	950714	1300	S	decane	1	0 ppbC in the midst of several days with concentrations above 1.4 ppbC
Lynn, MA	250092006	950720	1300	S	isoprene	1	0 ppbC between hours with concentrations above 5 ppbC
Lynn, MA	250092006	950720	1700	S	aceta	1	High concentration
Lynn, MA	250092006	950724	0800-2300	S	carbonyls	16	All carbonyl species and total because of missing C1-C2 species, high concentrations of others
Lynn, MA	250092006	950726	1300	S	NMOC, UNID	1	High unidentified concentrations
Lynn, MA	250092006	950809	1100	S	sample	1	Missing species
Lynn, MA	250092006	950826	0400	S	sample	1	High concentrations several species, anomalous benzene concentrations
Lynn, MA	250092006	950831	0500, 0600	S	sample	2	High concentrations several species, anomalous benzene concentrations

[a] Total number of hydrocarbon samples reported: 654 (June), 725 (July), 24 (August)
Total number of carbonyl samples reported: 183 (July)

Table A-4. Summary of suspect and invalid samples in the 1996 June-August Lynn, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments
Lynn, MA	250092006	960604	1300	S	sample	1	High 23dmp, 3mhxa
Lynn, MA	250092006	960709	2000	S	sample	1	Low concentrations
Lynn, MA	250092006	960702	1000, 1300	S	sample	2	High nundc
Lynn, MA	250092006	960802	0900	S	sample	1	High nundc
Lynn, MA	250092006	960827	1700	S	sample	1	High pdeben

Table A-5. Summary of suspect and invalid samples in the 1993 June-August Philadelphia, Pennsylvania PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Philadelphia, PA	421010004	930601	1100	S	sample	1	High 3m1be
Philadelphia, PA	421010004	930602	0500	S	sample	1	Misidentification problems
Philadelphia, PA	421010004	930606	2000	S	sample	1	High concentrations
Philadelphia, PA	421010004	930607	1100	S	sample	1	High 23dmb
Philadelphia, PA	421010004	930709	1400	S	sample	1	Misidentification problems
Philadelphia, PA	421010004	930721	2000	S	sample	1	Misidentification problems
Philadelphia, PA	421010004	930723	2000	S	sample	1	Misidentification problems
Philadelphia, PA	421010004	930825	0500, 2000, 2300	S	sample	3	2300 low concentration, looks like a blank
Philadelphia, PA	421010004	930809	1100	S	sample	1	High 3mheptane concentration
Philadelphia, PA	421010004	930702	1400	S	sample		High ispbz, 124tmb concentration
Philadelphia, PA	421010004	930616	0800, 1100, 1400	S	135tmb	3	High concentration
Philadelphia, PA	421010004	930618	2300	S	prpyl	1	High concentration
Philadelphia, PA	421010004	930828	2300	S	sample	1	Low concentration (looks like a blank)
Philadelphia, PA	421010004	930613	1100	S	3m1be	1	High concentration
Philadelphia, PA	421010004	930617	1400, 1700, 2000, 2300	S	3m1be	4	High concentration
Philadelphia, PA	421010004	930623	1400, 1700, 2000, 2300	S	3m1be	4	High concentration
Philadelphia, PA	421010004	930629	1100, 1400, 1700, 2000, 2300	S	3m1be	5	High concentration
Philadelphia, PA	421010004	930717	0200, 0500, 1100, 1400	S	3m1be	4	High concentration
Philadelphia, PA	421010004	930804	1400, 1700, 2000, 2300	S	3m1be	4	High concentration
Philadelphia, PA	421010004	930828	1700, 2000	S	3m1be	2	High concentration
Philadelphia, PA	421010004	930716	1400, 1700, 2000, 2300	S	3m1be	4	High concentration

[a] Other problems noted include cyclopentane, hexane, and 2,3-dimethylbutane misidentification during some time periods.

Table A-6. Summary of suspect and invalid samples in the 1994 June-August Philadelphia, Pennsylvania PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Philadelphia, PA	421010004	940607	2000	S	sample	1	High concentration
Philadelphia, PA	421010004	940729	0800	S	c2hex	1	High concentration
Philadelphia, PA	421010004	940823	0800	S	c2pne	1	High concentration
Philadelphia, PA	421010004	940606	1400, 1700	S	3m1be	2	High concentration
Philadelphia, PA	421010004	940811	1700	S	prpyl	1	High concentration
Philadelphia, PA	421010004	940618	1400, 1700, 2000, 1100	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940625	0200, 0500, 0800, 1100, 1400, 1700, 2000, 2300	S	3m1be	8	High concentration
Philadelphia, PA	421010004	940624	1100, 1400, 1700, 2000, 2300	S	3m1be	5	High concentration
Philadelphia, PA	421010004	940706	1100, 1400, 1700, 2000	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940712	1100, 1400, 1700, 2000	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940730	1100, 1400, 1700, 2000	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940805	1100, 1400, 1700, 2000	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940817	0800, 1400, 1700, 2000, 2300	S	3m1be	5	High concentration
Philadelphia, PA	421010004	940823	1400, 1700, 2000, 2300	S	3m1be	4	High concentration
Philadelphia, PA	421010004	940829	0800, 1400, 1700, 2000	S	3m1be	4	High concentration

[a] Other problems noted include t-2-pentene apparent detection limit of about 1 ppbC; 2-methyl-2-butene high when t-2-pentene is high.

Table A-7.   Summary of suspect and invalid samples in the 1995 June-August Philadelphia, Pennsylvania PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments
Philadelphia, PA	421010004	950601 to 950831	0000 to 2300	S	NMOC, unid	all	NMOC less than sum of the identified species, thus negative unidentified.
Philadelphia, PA	421010004	950601 to 950831	0000 to 2300	S	o-xylene	all	Constant concentration during the month at about 3.0 ppbC.
Philadelphia, PA	421010004	950620	0500	S	mcypnta	1	high concentration of this species compared to rest of data
Philadelphia, PA	421010004	950623	1400,1700	S	unid, NMOC	2	NMOC less than sum of the identified species, thus negative unidentified
Philadelphia, PA	421010004	950624	0800,1100	S	sample	2	possible cold trap failure based on low ethane concentrations
Philadelphia, PA	421010004	950626	1400	S	n-hexane	1	High concentration
Philadelphia, PA	421010004	950707	0200	S	propene	1	high concentration
Philadelphia, PA	421010004	950708	0200	S	propene	1	high concentration
Philadelphia, PA	421010004	950802	1400	S	NMOC, unid	1	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950803	0200, 1400, 1700	S	NMOC, unid	3	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950804	0200	S	propene	1	High concentration
Philadelphia, PA	421010004	950804	2300	S	NMOC, unid	1	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950805	1100, 1400	S	NMOC, unid	2	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950812	0800	S	NMOC, unid	1	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950813	0800, 1100, 1400, 1700, 2000	S	NMOC, unid	5	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950816	0800, 1100	S	NMOC, unid	2	NMOC less than sum of the identified species
Philadelphia, PA	421010004	950828	0500	S	aceta	1	High concentration
Philadelphia, PA	421010004	950814	2300	S	sample	1	Possible cold trap failure based on low ethane concentration

[a] Total number of hydrocarbon samples reported: 215 (June), 223 (July), 139 (August).

Table A-8. Summary of suspect and invalid samples in the 1996 June-August Philadelphia, Pennsylvania PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments
Philadelphia, PA	421010004	960726	0500	S	c2bte	1	High concentration
Philadelphia, PA	421010004	960714	0500	S	c2bte	1	High concentration
Philadelphia, PA	421010004	960713	0800	S	1pnte	1	High concentration
Philadelphia, PA	421010004	960712	0200	S	1pnte	1	High concentration
Philadelphia, PA	421010004	960708	0500	S	c2bte	4	High concentration
Philadelphia, PA	421010004	960619	0800	S	1pnte	1	High concentration
Philadelphia, PA	421010004	960617	0200, 2000	S	1pnte	2	High concentration
Philadelphia, PA	421010004	960616	1700	S	1pnte	1	High concentration
Philadelphia, PA	421010004	960808	1700	S	sample	1	ispnta concentration = 0
Philadelphia, PA	421010004	960809	1400	S	sample	1	High concentration
Philadelphia, PA	421010004	960821	1400, 2000	S	sample	2	High concentration
Philadelphia, PA	421010004	960822	2000	S	sample	1	High concentration
Philadelphia, PA	421010004	960823	1700, 2300	S	sample	2	High concentration
Philadelphia, PA	421010004	960827	1400, 1700, 2000	S	sample	3	High concentration
Philadelphia, PA	421010004	960820	1400	S	sample	1	High concentration
Philadelphia, PA	421010004	960829	1100	S	sample	3	High concentration
Philadelphia, PA	421010004	960824	1700	S	sample	1	High concentration

Table A-9.   Summary of suspect and invalid samples in the 1994, 1996 June-August, Chicopee, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Chicopee, MA	250130008	960711	1500	S	sample	1	High 22dmb
Chicopee, MA	250130008	960813	1000	S	sample	1	High pdeben
Chicopee, MA	250130008	960813	1200	S	sample	1	High nundc
Chicopee, MA	250130008	960618	0000	S	2mpna	1	Misidentification problems
Chicopee, MA	250130008	960617	0800, 1000, 1100, 1200	S	2mpna	4	Misidentification problems
Chicopee, MA	250130008	960608	0000, 0400	S	2mpna	2	Misidentification problems
Chicopee, MA	250130008	960610	1300	S	sample	1	Misidentification problems
Chicopee, MA	250130008	960614	0700	S	sample	1	Missing C2-C6
Chicopee, MA	250130008	960621	1500	S	sample	1	Missing C6+
Chicopee, MA	250130008	940614	1400-2200	S	sample	9	High n-undecane from calibration carryover
Chicopee, MA	250130008	940618	0300	S	sample	1	Poor C2-C6 results
Chicopee, MA	250130008	940615	1000-1600	S	sample	9	Misidentification problems
Chicopee, MA	250130008	940618	0400, 0600, 0700	S	sample	3	Missing C2-C6
Chicopee, MA	250130008	940810	0800-1300	S	sample	6	Missing key species
Chicopee, MA	250130008	940708	1300-1500	S	sample	3	Missing key species
Chicopee, MA	250130008	940707	1500	S	sample	1	Missing key species
Chicopee, MA	250130008	940722	1300	S	sample	1	High cypnta

1994: Possible problems with acetylene, xylenes identification and n-undecane carryover from calibrations. Some missing NMHC and other data in August. Cyclopentane concentrations appeared high during the daytime.

1996: High n-hexane throughout the year.

Table A-10. Summary of suspect and invalid samples in the 1995 June-August Chicopee, Massachusetts PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments
Chicopee, MA	250130008	950607 to 950608	1400 to 1000	S	sample	20	Odd relative proportions of species in these samples
Chicopee, MA	250130008	950626	0700	S	sample	1	Low n-pentane, other odd relative proportions of species
Chicopee, MA	250130008	950702	0900	S	c2-butene	1	High concentration
Chicopee, MA	250130008	950706	0700	S	UNID	1	High unidentified concentration
Chicopee, MA	250130008	950706	2000, 2300	S	forma, acet, aceta	2	Very high concentrations, e.g., above 100 ppbC formaldehyde
Chicopee, MA	250130008	950707	0200, 0500	S	forma, acet, aceta	2	Very high concentrations, e.g., above 100 ppbC formaldehyde
Chicopee, MA	250130008	950708	0700	S	UNID	1	High unidentified concentration
Chicopee, MA	250130008	950717	1100-2300	S	UNID	13	High unidentified concentration
Chicopee, MA	250130008	950717	1800	S	n-decane	1	High concentration
Chicopee, MA	250130008	950717	1300,1400, 1600,1700, 1900-2100,	S	2m1pte	7	High concentration
Chicopee, MA	250130008	950718	0500	S	2m1pte	1	High concentration
Chicopee, MA	250130008	950720	1700	S	acet	1	High concentration
Chicopee, MA	250130008	950721	2000	S	forma, acet, aceta	1	Very high concentrations, e.g., above 100 ppbC formaldehyde
Chicopee, MA	250130008	950807	0700	S	sample	1	High unidentified
Chicopee, MA	250130008	950819	0700	S	sample	1	Low alkanes, high unidentified
Chicopee, MA	250130008	950725	2000	S	forma, acet, aceta	1	Very high concentrations, e.g., above 100 ppbC formaldehyde

[a] Total number of hydrocarbon samples reported: 574 (June), 388 (July), 476 (August)
    Total number of carbonyl samples reported: 50 (June)

Table A-11. Summary of suspect and invalid hydrocarbon samples in the 1994 June-August East Hartford, CT PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments
East Hartford, CT	090031003	940614	1100	I	sample	1	Calibration
East Hartford, CT	090031003	940720	1700-2300	S	sample	7	Missing C6+ data
East Hartford, CT	090031003	940721	0000-2300	S	sample	24	Missing C6+ data
East Hartford, CT	090031003	940723	0000-2300	S	sample	24	Missing C6+ data
East Hartford, CT	090031003	940715	0000-2300	S	sample	24	Misidentification problems
East Hartford, CT	090031003	940622	0800	S	sample	1	Calibration carryover
East Hartford, CT	090031003	940708	1400	S	sample	1	Missing data
East Hartford, CT	090031003	940610	0700	S	sample	1	Missing data
East Hartford, CT	090031003	940724	0000, 0100	S	sample	2	Missing C6+ data
East Hartford, CT	090031003	940722	0000-2300	S	sample	24	Missing C6+ data

Table A-12. Summary of suspect and invalid hydrocarbon samples in the 1995 June-August East Hartford, CT PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments
East Hartford, CT	090031003	950524	1300-1800	S	sample	6	Anomalous high concentrations of 1-pentene, several other species.
East Hartford, CT	090031003	950612	1400-1900	S	UNID, NMOC	6	Period of extremely high concentration unidentified.
East Hartford, CT	090031003	950614	1100	S	UNID, NMOC	1	Period of extremely high concentration unidentified.
East Hartford, CT	090031003	950701 to 950731	0000 to 2300	I	all hydrocarbon data	744	Invalidated by State of CT due to instrument problems. Carbonyl data not affected.
East Hartford, CT	090031003	950716	0500	S	sample	1	Anomalous low concentration carbonyls.
East Hartford, CT	090031003	950722	0200	S	sample	1	Anomalous low concentration carbonyls.

[a] Total number of hydrocarbon samples reported: 88 (May), 402 (June), 744 (July);
    Total number of carbonyl samples reported: 127 (July)

Table A-13. Summary of suspect and invalid samples in the 1993-1996 Houston (Clinton Drive), Texas PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments
Houston, TX	482011035	930711	1700	S	sample	1	High C2-C4; low C5+
Houston, TX	482011035	930626	2000	S	sample	1	Odd fingerprint
Houston, TX	482011035	930712	1400	S	sample	1	High C7+
Houston, TX	482011035	930830	0300	S	sample	1	High C4-C7
Houston, TX	482011035	930827	1900	S	sample	1	High hexane
Houston, TX	482011035	930716	1700	S	sample	1	Odd fingerprint
Houston, TX	482011035	930805	0800, 1200	S	ebenz, styr	2	High concentration
Houston, TX	482011035	930721	0700, 1300, 1400, 1800, 2200	S	135tmb	5	High concentration
Houston, TX	482011035	930722	0200	S	135tmb	1	High concentration
Houston, TX	482011035	930807	0200	S	sample	1	High 124tmb
Houston, TX	482011035	930829	0500	S	bpine	1	High concentration
Houston, TX	482011035	930830	2200	S	bpine	1	High concentration
Houston, TX	482011035	940913	0000	S	cypne	1	High concentration
Houston, TX	482011035	950713	1600	S	sample	1	High 2m1pe
Houston, TX	482011035	950724	1500	S	sample	1	Calibration data
Houston, TX	482011035	950725	0200	S	sample	1	High concentrations
Houston, TX	482011035	950629	0800-2200	S	sample	16	Missing C6+ species
Houston, TX	482011035	950630	0800-1300	S	sample	6	Missing C6+ species
Houston, TX	482011035	950601	1500	S	sample	1	Missing C6+ species
Houston, TX	482011035	950606	1600	S	sample	1	Missing C6+ species
Houston, TX	482011035	950607	1000	S	sample	1	Missing C6+ species
Houston, TX	482011035	960829	2300	S	sample	1	Odd fingerprint
Houston, TX	482011035	960822	0300	S	sample	1	Odd fingerprint
Houston, TX	482011035	960827	1400	S	sample	1	Missing key species
Houston, TX	482011035	950713	2100	S	sample	1	High concentrations

Table A-14. Summary of suspect and invalid samples in the 1995 June-August Bronx, New York PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments[b]
Bronx, NY	360050083	950519	0400	S	sample	1	Looks like calibration gas carryover into this sample (n-nonane to n-undecane).
Bronx, NY	360050083	950713 to 950721	1900 to 0800	I	alkene, alkane totals	182	Speciated data eluting past 3-mpnta were not reported, NMOC not reported, no unidentified calculated.
Bronx, NY	360050083	950714	0200	S	isoprene	1	High concentration for this time of day
Bronx, NY	360050083	950821 to 950829	0900 to 1600	S	samples	184	Missing benzene, toluene, and other species concentration data
Bronx, NY	360050083	950907	1300	S	sample	1	Very high concentration (15 ppmC)

[a] Total number of hydrocarbon samples reported: 386 (May), 566 (July), 695 (August), 278 (September).
[b] 1994 data, no NMHC, no unidentified, 3-hr, every third day. There appear to be problems with acetylene.
  1996 data, 2, 3-dimethylbutane odd diurnal profiles, possible misidentification.

Table A-15.   Summary of suspect and invalid samples in the 1994-1996 June-August East Providence, RI PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples[a]	Comments
E. Providence, RI	440071010	950614	0500	S	benzene	1	Very high concentration of this species only
E. Providence, RI	440071010	950720	0500	S	sample	1	Ethane concentration = 0, suspect GC problem
E. Providence, RI	440071010	950814	0500	I	sample	1	calibration
E. Providence, RI	440071010	950815	0500	S	sample	1	High acetylene
E. Providence, RI	440071010	940619	0900	I	sample	1	all zero values
E. Providence, RI	440071010	940827	1800	I	sample	1	all zero values
E. Providence, RI	440071010	940619	1200	S	sample	1	missing C2's
E. Providence, RI	440071010	940620	1800, 2100	S	samples	2	missing C2's
E. Providence, RI	440071010	940621	0900-1800	S	samples	4	missing C2's
E. Providence, RI	440071010	940622	1800	S	sample	1	missing C2's
E. Providence, RI	440071010	940701	0600, 1500, 1800	S	sample	3	missing C2's
E. Providence, RI	440071010	940702	0900, 1200	S	sample	2	missing C2's
E. Providence, RI	440071010	940707	0000, 0900	S	sample	2	missing C2's
E. Providence, RI	440071010	940708	0900, 1500	S	sample	2	missing C2's
E. Providence, RI	440071010	940710	0900, 1500, 1800, 2100	S	sample	4	missing C2's
E. Providence, RI	440071010	940711	0000-1500, 2100	S	sample	7	missing C2's
E. Providence, RI	440071010	940712	0000-1500, 2100	S	sample	7	missing C2's
E. Providence, RI	440071010	940713	1500, 2100	S	sample	2	missing C2's
E. Providence, RI	440071010	940714	0000-1200	S	sample	5	missing C2's
E. Providence, RI	440071010	960730	2300	I	sample	1	all zero values
E. Providence, RI	440071010	960727	2300	I	sample	1	all zero values
E. Providence, RI	440071010	960620	2300	I	sample	1	all zero values
E. Providence, RI	440071010	960722	1100	S	sample	1	Ethane concentration = 0, suspect GC problems
E. Providence, RI	440071010	960728	1700	S	sample	1	Ethane concentration = 0, suspect GC problems
E. Providence, RI	440071010	960602	0500	S	sample	1	Ethane concentration = 0, suspect GC problems

[a]   Total number of hydrocarbon samples reported: 386 (May), 566 (July), 695 (August), 278 (September).

Table A-16.   Summary of suspect and invalid samples in the 1993 June-August Baton Rouge, Louisiana PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Baton Rouge, LA	220330008	930818	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930819	0900	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930713	0000	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930822	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930827	0600, 2100	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930821	0900	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930830	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930721	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930815	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930828	0300, 0600, 0900, 1200	S	NMHC	4	NMHC less than sum of species
Baton Rouge, LA	220330008	930729	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930726	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930816	0300	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930818	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930601	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930607	1500, 2100	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930613	1200	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930615	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930630	1500	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930706	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930710	1500	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930716	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930727	0900	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930803	0900	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930805	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930806	0000, 0300, 0600	S	NMHC	3	NMHC less than sum of species
Baton Rouge, LA	220330008	930807	2100	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930808	0000	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930810	0000, 0900	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930811	0300, 0600	S	NMHC	2	NMHC less than sum of species

Table A-16. Summary of suspect and invalid samples in the 1993 June-August Baton Rouge, Louisiana PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Baton Rouge, LA	220330008	930812	2100, 0000	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930814	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930813	0000, 2100	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930815	0900, 1200	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930816	0900, 1800, 2100	S	NMHC	3	NMHC less than sum of species
Baton Rouge, LA	220330008	930819	0300, 0600, 2100	S	NMHC	3	NMHC less than sum of species
Baton Rouge, LA	220330008	930820	0000, 0300, 0600, 1200	S	NMHC	4	NMHC less than sum of species
Baton Rouge, LA	220330008	930825	0600, 2100	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930824	0600, 1500	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930818	0600	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930822	0000, 0300, 0600	S	NMHC	3	NMHC less than sum of species
Baton Rouge, LA	220330008	930813	0000	S	sample	1	High ispbz (781 ppbC)
Baton Rouge, LA	220330008	930821	0300, 0600	S	NMHC	2	NMHC less than sum of species
Baton Rouge, LA	220330008	930813	0000	S	sample	1	High concentrations
Baton Rouge, LA	220330008	930812	0000	S	sample	1	High 135tmb
Baton Rouge, LA	220330008	930717	0000	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	930722	1500, 1800	S	NMHC	2	NMHC less than sum of species

[a] 1993: problems with NMHC values; acetylene appears low and erratic.

Table A-17. Summary of suspect and invalid samples in the 1994-1996 June-August Baton Rouge, LA PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments[a]
Baton Rouge, LA	220330008	940811	0000	S	NMHC	1	NMHC less than sum of species
Baton Rouge, LA	220330008	940620	1200, 1500	S	sample	2	Zero values important species
Baton Rouge, LA	220330008	940606	0900	S	sample	1	Very high concentrations
Baton Rouge, LA	220330008	940601	1200	S	sample	1	Zero values important species
Baton Rouge, LA	220330008	940605	1800	S	sample	1	Zero values important species
Baton Rouge, LA	220330008	940730	0600	S	sample	1	Misidentification problems
Baton Rouge, LA	220330008	940808	0600	S	sample	1	Misidentification problems
Baton Rouge, LA	220330008	940812	0600	S	sample	1	Misidentification problems
Baton Rouge, LA	220330008	950602	0000, 0300	S	sample	2	High concentrations
Baton Rouge, LA	220330008	950604	2100	S	sample	1	Missing data
Baton Rouge, LA	220330008	950605	1800	S	sample	1	High C7+ values
Baton Rouge, LA	220330008	950619	1800	S	sample	1	High 123tmb
Baton Rouge, LA	220330008	950623	0600	S	sample	1	Missing data
Baton Rouge, LA	220330008	960629	0600	S	sample	1	Missing data

[a] 1994: Ethane and ethene concentrations were not well-resolved due to humidity problems with the GC column and the effects of humidity on retention time. isopentane appears to have been incorrectly identified as 3-methyl-1-butene during much of the summer. The species 2-methylpentane appears to have a coeluter as evidenced by time series anomalies and its relationship to 3-methylpentane. Stoeckenius et al., also noted a possible coeluter with 1,2,4-trimethylbenzene, however, inspection of time series and scatter plots did not reveal an obvious problem.

Table A-18. Summary of suspect and invalid samples in the 1994-1995 June-August Harrington Beach, WI PAMS database.

Site, State	AIRS Code	Date	Time (ST)	QC	Species or Sample	Flagged Samples	Comments
Harrington Beach, WI	550890009	940710	1500	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	940713	1000, 1500	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	940808	1500	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	940809	0500, 1000, 1500	S	NMHC	3	NMHC less than sum of species
Harrington Beach, WI	550890009	940815	1000, 1500	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	940728	1000, 1500	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	940821	1000	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950828	1100	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950825	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950819	1100	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950812	0500, 1400	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	950807	1100	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950804	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950801	0500	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950729	1100	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950723	1100, 1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950717	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950708	1400, 1100	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	950702	1100, 1400	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	950629	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950614	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	940611	1500, 1100, 0500	S	NMHC	3	NMHC less than sum of species
Harrington Beach, WI	550890009	950602	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950608	0500, 1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950607	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950801	1100	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950729	1400	S	NMHC	1	NMHC less than sum of species
Harrington Beach, WI	550890009	950822	1100, 1400	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	950831	1100, 1400	S	NMHC	2	NMHC less than sum of species
Harrington Beach, WI	550890009	950720	1100, 1400	S	NMHC	2	NMHC less than sum of species

APPENDIX B

STATISTICAL TEST RESULTS FROM COMPARISONS OF ANNUAL DATA

The following two tables contain p-values from the Kruskal-Wallis nonparametric tests comparing 1994 to 1995 (Table B-1) and 1995 to 1996 (Table B-2) data sets for several species concentrations, weight fractions, and ratios. The data were also compared using a t-test. Statistical significance was assumed for p-values less than 0.05. Typically, the p-values were similar for both tests. When the p-values were in disagreement (i.e., one below 0.05 and one above), we have provided both values in the table (p-value from Kruskal-Wallis nonparametric test/p-value from t-test).

Table B-1. The p-values from comparisons of 1994 and 1995 data using the Kruskal-Wallis nonparametric tests. When the t-test and nonparametric test results did not agree, the p-values are given as follows: p-value from the nonparametric test / p-value from the t-test.

Species, Ratios	Los Angeles, CA	Chicago, IL	New York City (Bronx)	Springfield MA (Chicopee)	E. Hartford, CT	Boston, MA (Lynn)	Washington DC (McMillan Reservoir)	Philadelphia, PA	Milwaukee, WI	E. Providence, RI	Harrington Beach, WI	Baton Rouge, LA
Concentrations:												
n-Butane	0.6	0.001	0.12	0.04	0.25	<0.001	<0.001	<0.001	<0.001	0.002/0.07	0.04	0.06
1,3-Butadience	0.75	na	na	na	na	na	na	na	na	na	na	na
Benzene	0.52	0.03	0.003	<0.001	0.02	<0.001	<0.001	<0.001	<0.001	0.11	<0.001	0.05/0.04
Xylenes	0.22	<0.001	0.16	<0.001	0.01/0.16	0.02	<0.001	0.001	0.002	0.63	0.001	0.5
Formaldehyde	na	0.49	na	0.008/0.1	0.02/0.1	0.001	na	0.08	na	0.08	na	na
i-Butene	<0.001	na	na	na	na	na	na	na	na	na	na	na
Trimethylbenzenes	0.28	0.006	0.35	<0.001	0.03/0.4	0.02	<0.001	<0.001	0.002	0.94	0.02	0.42
TNMOC	0.02	na	na	0.8	0.03/0.05	<0.001	0.3	0.002	<0.001	0.3	<0.001	<0.001
Isoprene	0.04	0.06/0.02	0.06	0.01	0.06	0.83	0.11/<0.001	0.17	0.06/0.045	0.24	0.85	0.02/0.07
Weight Percent:												
n-Butane	0.002	na	na	<0.001	0.13	<0.001	<0.001	<0.001	0.02	<0.001	0.18	<0.001
1,3-Butadience	0.26	na	na	na	na	na	na	na	na	na	na	na
Benzene	<0.001	na	na	<0.001	0.4	<0.001	<0.001	<0.001	0.01	0.001/0.4	0.03/0.1	<0.001
Xylenes	0.005	na	na	<0.001	0.21	0.43	<0.001	0.23	0.68	0.67	0.46	<0.001
i-Butene	<0.001	na	na	na	na	na	na	na	na	na	na	na
Trimethylbenzenes	0.06/0.03	na	na	<0.001	0.02/0.6	0.07	<0.001	0.27	na	0.19	0.66	<0.001
Isoprene	0.34	na	na	<0.001	<0.001	0.19	0.24	<0.001	0.96	0.92	0.32	0.45/0.02
Ratios:												
Benzene/Toluene	<0.001	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.94	<0.001/0.052	0.12	<0.001
n-Butane/i-Pentane	0.003	0.02/0.6	0.03	0.06/0.03	<0.001	<0.001	<0.001	<0.001	0.001	<0.001	0.11	0.001
Benzene/Acetylene	<0.001	0.01	0.006/0.7	0.1	<0.001	<0.001	<0.001	<0.001	<0.001	0.001/0.3	<0.001	<0.001
TMB/Toluene	0.29	<0.001	0.76	0.02	0.12	0.18	<0.001	0.04	0.06	0.82	0.38	0.29
TMB/Xylenes	0.67	0.001	0.03	0.001	0.27	0.51	<0.001	0.04	0.71	0.03	0.27	0.48

na = data inadequate for the analysis (e.g., missing NMHC, not analyzed for, etc.)

Table B-2.   The p-values from comparisons of 1995 and 1996 data using the Kruskal-Wallis nonparametric tests. When the t-test and nonparametric test results did not agree, the p-values are given as follows: p-value from the nonparametric test / p-value from the t-test.

Species, Ratios	Los Angeles, CA	Chicago, IL	New York City (Bronx)	Springfield MA (Chicopee)	E. Hartford, CT	Boston, MA (Lynn)	Wasahington DC (McMillan Reservoir)	Philadelphia, PA	Milwaukee, WI	E. Providence, RI	Harrington Beach, WI	Baton Rouge, LA
Concentrations:												
n-Butane	0.14	0.08	<0.001	0.37	n	0.1	<0.001	0.76	<0.001	0.01	0.45	0.003
1,3-Butadience	0.41	na	na	na	o	na	na	na	0.18	na	0.44	na
Benzene	0.04	<0.001	0.02	0.06/0.03		0.69	0.046/0.1	0.01/0.07	<0.001	<0.001	0.61	0.46
Xylenes	0.2	<0.001	0.37/0.01	0.15/0.02	d	<0.001	<0.001	0.06	<0.001	<0.001	0.97	0.45
Formaldehyde	na	<0.001	0.32	<0.001/0.08	a	0.16	0.035	0.39	<0.001	<0.001	0.006	<0.001
i-Butene	0.79	na	na	na	t	na	na	na	na	na	na	na
Trimethylbenzenes	0.12	<0.001	0.37	0.001	a	<0.001	<0.001	0.002	0.001	<0.001	0.76	<0.001
TNMOC	0.12	<0.001	0.004	<0.001		<0.001	0.28/0.01	0.08	0.99	0.009	0.003	0.002
Isoprene	0.1	0.004/0.19	0.56	0.19		<0.001	<0.001	0.01/0.75	0.02	<0.001	0.08	0.88
Weight Percent:												
n-Butane	0.23	<0.001	<0.001	0.02		<0.001	<0.001	0.03/0.10	0.72	0.24	0.01	<0.001
1,3-Butadience	0.1	na	na	na		na	na	na	0.15	na	0.42	na
Benzene	<0.001	0.002	<0.001	<0.001		<0.001	0.5/0.047	0.09	0.006	<0.001	<0.001	<0.001
Xylenes	0.37	0.63	0.56	<0.001		0.2	<0.001	0.01	0.045/0.16	<0.001	0.13	0.003
i-Butene	0.06	na	na	na		na	na	na	na	na	na	na
Trimethylbenzenes	0.13	0.92/<0.001	<0.001	<0.001		0.2	<0.001	0.11	0.28	<0.001	0.94	0.12
Isoprene	0.66	0.01/0.25	0.62	0.29/0.04		0.1	<0.001	<0.001/0.27	0.41	0.03/0.13	0.17	0.005
Ratios:												
Benzene/Toluene	<0.001	<0.001	0.02/0.19	<0.001/0.16		<0.001	0.28	0.13	0.005	0.02/0.9	0.19	0.28
n-Butane/i-Pentane	0.005	0.002	0.001	<0.001		0.49	<0.001	<0.001	<0.001	0.61	0.26	0.01/0.45
Benzene/Acetylene	<0.001	0.003	<0.001	0.002/0.99		<0.001/0.3	<0.001	0.07	<0.001	0.08	0.98	0.003/0.57
TMB/Toluene	0.19	0.002	<0.001	0.19/0.03		0.003	<0.001	<0.001	<0.001	0.004	0.97	<0.001
TMB/Xylenes	0.91	0.06/<0.001	<0.001	<0.001		0.004/0.3	<0.001	<0.001	<0.001	<0.001	0.89	<0.001

na = data inadequate for the analysis (e.g., missing NMHC, not analyzed for, etc.)