IEc

Section 812 Prospective Study of the Benefits and Costs of the Clean Air Act:

Air Toxics Case Study - Health Benefits of Benzene Reductions in Houston, 1990-2020

Final Report | July 14, 2009

Prepared for:

Office of Policy Analysis and Review

U.S. Environmental Protection Agency

Washington, DC 20460

Prepared by:

Industrial Economics, Incorporated

2067 Massachusetts Avenue

Cambridge, MA 02140

617/354-0074

INDUSTRIAL ECONOMICS, INCORPORATED

TABLE OF CONTENTS

LIST OF ACRONYMS

EXECUTIVE SUMMARY

CHAPTER 1: INTRODUCTION

CHAPTER 2: ANALYTIC APPROACH

CHAPTER 3: RESULTS

CHAPTER 4: DISCUSSION AND CONCLUSIONS

REFERENCES

APPENDICES

APPENDIX A: BENZENE EMISSIONS MODELING REPORT

APPENDIX B: BENZENE AIR QUALITY AND EXPOSURE MODELING REPORT

APPENDIX C: BENZENE HEALTH EFFECTS LITERATURE REVIEW

APPENDIX D: LIFE-TABLE MODEL EQUATIONS

APPENDIX E: ATTACHED GARAGE ANALYSIS EQUATIONS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AEO	Annual Energy Outlook
AERMET	AERMOD Meteorological Preprocessor
AERMOD	American Meteorological Society/Environmental Protection Agency Regulatory Model
ALL	Acute Lymphocytic Leukemia
AML	Acute Myelogenous Leukemia
ANLL	Acute Non-lymphocytic Leukemia
AQM	Air Quality Modeling
ATP	Anti-Tampering Program
BEIR	Biological Effects of Ionizing Radiation
CAA	Clean Air Act
CAAA	Clean Air Act Amendments of 1990
CalEPA	California Environmental Protection Agency
CAMD	Clean Air Markets Division
CHAD	Consolidated Human Activity Database
CI	Compression Ignition
CLL	Chronic Lymphocytic Leukemia
CMAQ	Community Multi-scale Air Quality
CML	Chronic Myelogenous Leukemia
COI	Cost-of-Illness
DOE	U.S. Department of Energy
EC	Exposure Concentration
EGU	Electricity Generating Unit
EIA	U.S. Department of Energy’s Energy Information Administration
EMS-HAP	Emissions Modeling System for Hazardous Air Pollutants
EPA	U.S. Environmental Protection Agency

ESD	Emission Standards Division
HAP	Hazardous Air Pollutant
HAPEM	Hazardous Air Pollutant Exposure Model
HL	Hodgkin's Lymphoma
I/M	Inspection and Maintenance
IEc	Industrial Economics, Incorporated
IRIS	Integrated Risk Information System
IUR	Inhalation Unit Risk
MACT	Maximum Achievable Control Technology
ME	Microenvironment
MOA	Mode of Action
MSAT	Mobile Source Air Toxics
NAAQS	National Ambient Air Quality Standards
NATA	National Air Toxics Assessment
NEI	National Emissions Inventory
NESHAP	National Emissions Standards for Hazardous Air Pollutants
NHL	Non-Hodgkin's Lymphoma
NLEV	National Low-Emission Vehicle
NMIM	National Mobile Inventory Model
NPV	Net Present Value
NRC	National Research Council
NWS	National Weather Service
OAR	Office of Air and Radiation
OGWDW	Office of Ground Water and Drinking Water's
OTAQ	Office of Transportation and Quality
PM	Particulate Matter
POTW	Publicly Owned Treatment Works
RfC	Reference Concentration
RFG	Reformulated Gasoline
RIA	Regulatory Impact Analysis
RR	Relative Risk

RVP	Reduced Vapor Pressure
SAB	Science Advisory Board
SAB Council	Science Advisory Board Advisory Council for Clean Air Compliance Analysis
SAB EEAC	Science Advisory Board Environmental Economics Advisory Committee
SAB HES	Science Advisory Board (SAB) Health Effects Subcommittee
SCCs	Source Classification Codes
SI	Spark Ignition
SOCMI HON	Synthetic Organic Chemical Manufacturing Industry Hazardous Organic NESHAP
TCEQ	Texas Council on Environmental Quality
TPY	Tons Per Year
TRI	Toxics Release Inventory
TTI	Texas Transportation Institute
VMT	Vehicle Miles Traveled
VOC	Volatile Organic Compound
VSL	Value of Statistical Life
WSC	West South Central
WTP	Willingness to Pay

EXECUTIVE SUMMARY

Section 812 of the Clean Air Act Amendments of 1990 (CAAA) requires the U.S. Environmental Protection Agency (EPA) to perform periodic, comprehensive analyses of the total costs and total benefits of programs implemented pursuant to the Clean Air Act (CAA). EPA has completed two of these analyses: a retrospective analysis in 1997 of the original CAA covering the period 1970 to 1990, and a prospective analysis in 1999 of the incremental costs and benefits of the CAAA over the period 1990 to 2010. In both of these studies, estimation of the benefits of reduced concentrations of hazardous air pollutants (HAPs) has proved difficult, due to gaps in the toxicological database; difficulty in designing population-based epidemiological studies with sufficient power to detect health effects; limited ambient and personal exposure monitoring data; limited data to estimate exposures in some critical microenvironments; and insufficient economic research to support valuation of the types of health impacts often associated with exposure to individual HAPs.

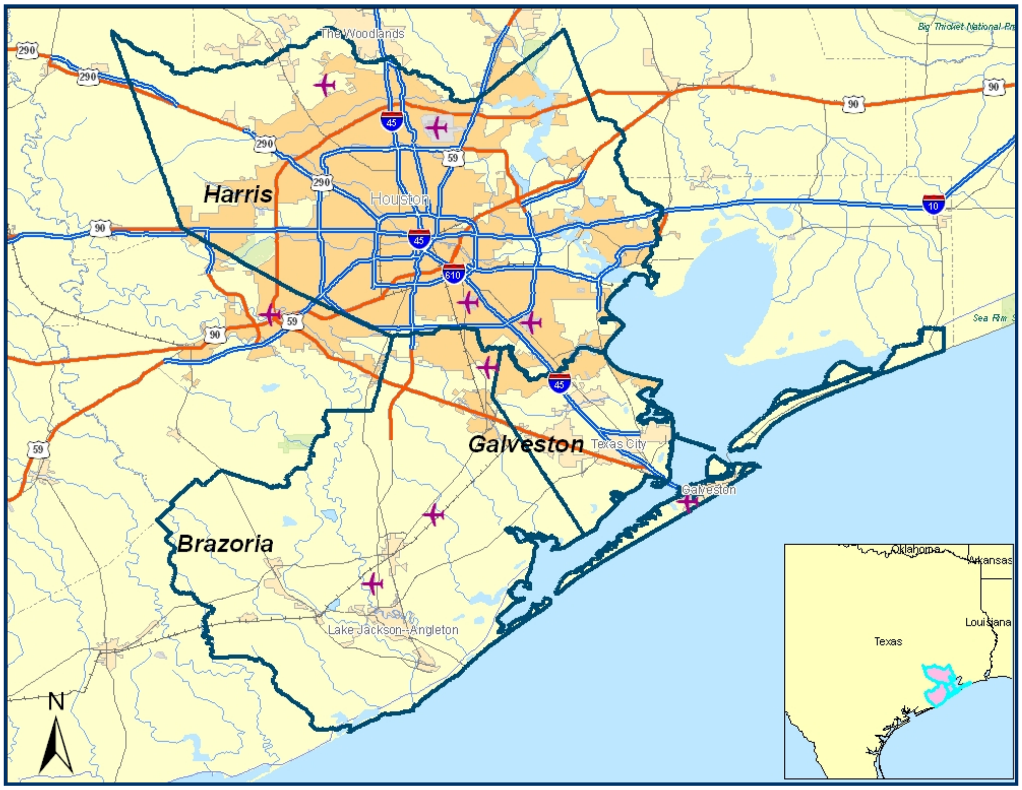
In 2001, EPA's Science Advisory Board Advisory Council for Clean Air Compliance Analysis (SAB) proposed that EPA undertake a HAP benefits case study of a well-studied HAP such as benzene to accompany EPA’s second prospective cost-benefit analysis of the CAAA. The SAB indicated that such a study should identify limitations and data gaps; provide an estimate of uncertainties; and provide a scientific basis for deciding whether there is merit in pursuing a greater ability to address air toxics. In response to these comments, EPA developed a methodology for estimating the health benefits of benzene reductions and has applied it in a metropolitan-scale case study of the benefits of CAAA controls on benzene emissions to accompany the main 812 analysis. The results of this study are described in this report.

This case study has two main objectives. The first is to demonstrate a methodology that generates human health benefits resulting from CAAA controls on a single HAP in an urban setting, while highlighting key limitations and uncertainties. The second is to provide a basis for considering more broadly the value of such an exercise for HAP benefits characterization nationwide. This case study is not intended to provide a comprehensive assessment of the benefits of benzene reductions due to the Clean Air Act.

We selected the Houston-Galveston area for the case study (Figure ES-1), a metropolitan area with a large population (a total of 3.4 million in 2000, with nearly 3 million people in Harris County alone) and significant benzene emissions from both on-road mobile sources and large industrial point sources such as petroleum refineries. The study area includes Harris, Galveston, and Brazoria counties – the three counties responsible for

99% of the point source emissions in Houston metropolitan area, according to EPA's 1999 National Emissions Inventory (NEI).

FIGURE ES-1: BENZENE CASE STUDY AREA



The timeframe for this analysis, 1990 through 2020, matches that used in the criteria pollutant analysis of the second prospective Section 812 study. In addition to the base year, 1990, we model results for three target years, 2000, 2010, and 2020.

We conducted this benefits analysis using the standard approach applied in the main 812 criteria pollutant analysis, which includes the following five steps:

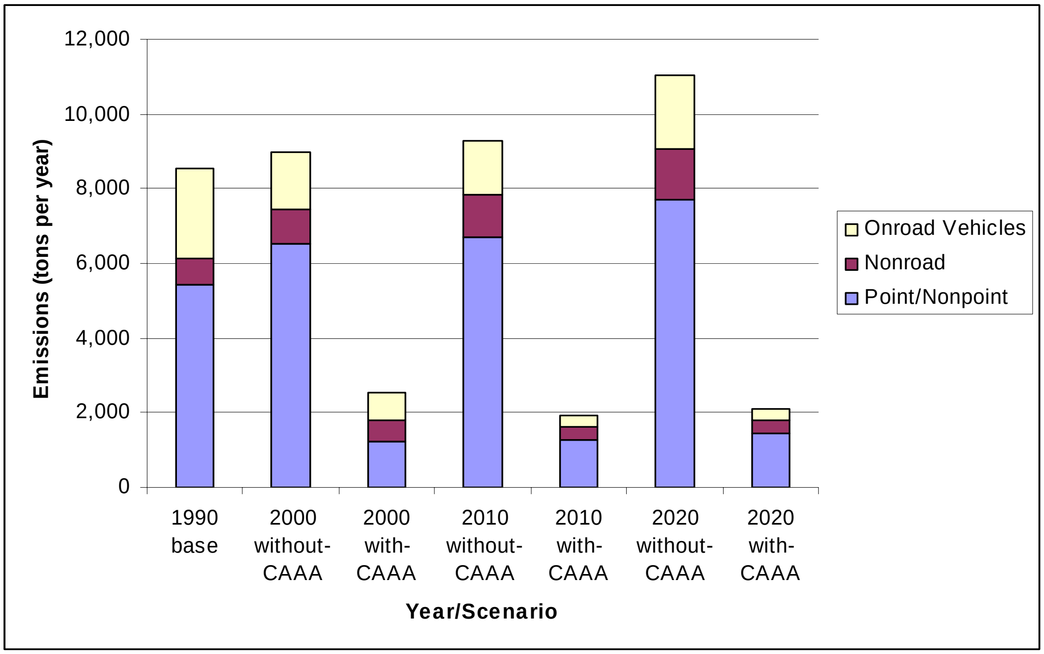
1. Scenario Development
2. Emissions Estimation
3. Air Quality and Exposure Modeling
4. Health Effects Modeling
5. Valuation

We model benzene exposures and health impacts under two scenarios, one reflecting the impacts of all regulatory programs affecting benzene that were enacted in response to the 1990 CAAA (the *With-CAAA* scenario), and one assuming no additional benzene pollution control activity beyond the regulatory requirements existing in 1990 (the

Without-CAAA scenario).[1] The difference between the two scenarios reflects the impact of the CAAA on benzene concentrations and benzene-related health effects in the study area.

We estimated benzene emissions in the Houston-Galveston study area for four source categories: point, non-point (formerly "area sources"), on-road, and non-road. Exhibit ES-2 illustrates emissions changes in each category due to CAAA programs, with significant reductions observed in all categories compared to the *Without-CAAA* case.

FIGURE ES-2: MAJOR, AREA AND OTHER, ON-ROAD, AND NON-ROAD EMISSIONS (TONS) FOR EACH YEAR AND SOURCE TYPE



We applied EPA's American Meteorological Society/U.S. EPA Regulatory Model (AERMOD) dispersion modeling system (U.S. EPA 2004b) to convert emissions estimates to ambient benzene concentrations in the Houston-Galveston study area. Following completion of the AERMOD runs, we applied EPA's Hazardous Air Pollutant Exposure Model, Version 6 (HAPEM6) to the hourly ambient benzene concentration output from AERMOD to generate time-weighted average benzene exposure concentrations for the study population. The HAPEM results reflect the average benzene concentrations likely to be experienced by the study population as they carry out their daily activities.

---

[1] Our modeling does not include indoor sources of exposure.

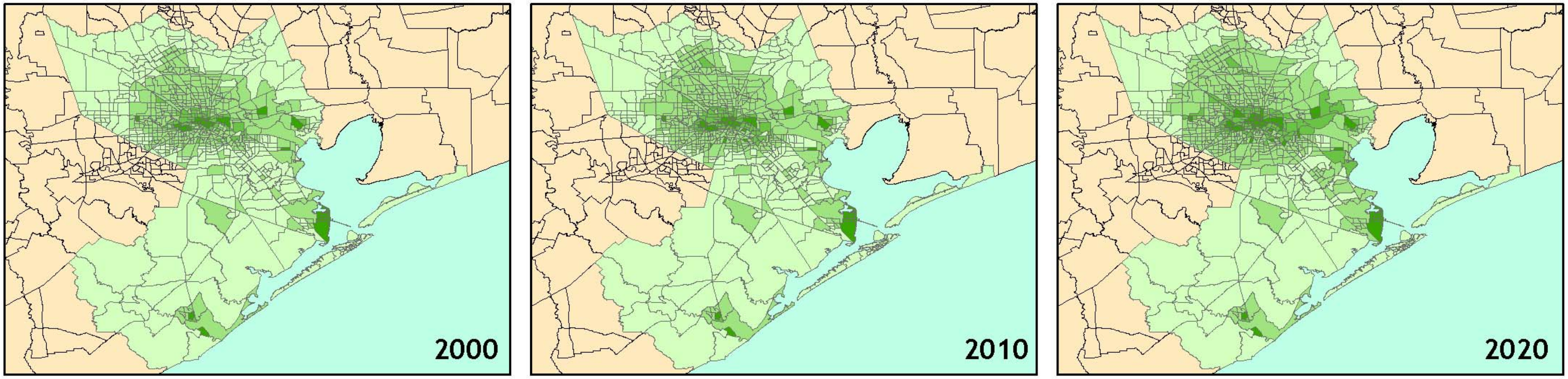
Figure ES-3 presents maps showing the spatial distribution of benzene reductions across the study area. The top row of maps shows the AERMOD estimates of the reduction in annual average ambient benzene levels due to CAAA programs in 2000, 2010, and 2020. The bottom row shows the same progression using the exposure concentration results from the HAPEM model. The maps show the greatest reductions (in excess of 5 $\mu g/m^3$) occur in Harris County in the downtown Houston area, within the rings of the interstate; in the Texas City area of Galveston County where a number of refineries and chemical facilities are located; and in southeastern Brazoria County, which also features major chemical manufacturing and petroleum refining facilities. Mobile source emission controls are a significant contributor to the reductions in Harris County, and we observe an increase over time in the extent and magnitude of reductions in that area, as mobile source controls become more effective over time. In general, HAPEM tends to smooth and spread out the AERMOD concentration changes; this reflects both aggregating results to the census tract level and incorporating the impact of commuting and other activities on the concentration experienced by the population in each census tract.

We focused our health benefits analysis on quantifying avoided cases of leukemia (all types), based on an extensive review of the available health effects literature. To estimate the avoided cases associated with benzene reductions in the study area, we constructed a life-table based risk assessment model. The life-table model assessed age-specific risks within each census tract in each year of the study, based on county-specific background rates of leukemia mortality and morbidity, age-specific benzene exposure data generated by HAPEM (and interpolated for unmodeled years) and a dose-response function from Crump (1994) relating benzene exposure with leukemia.

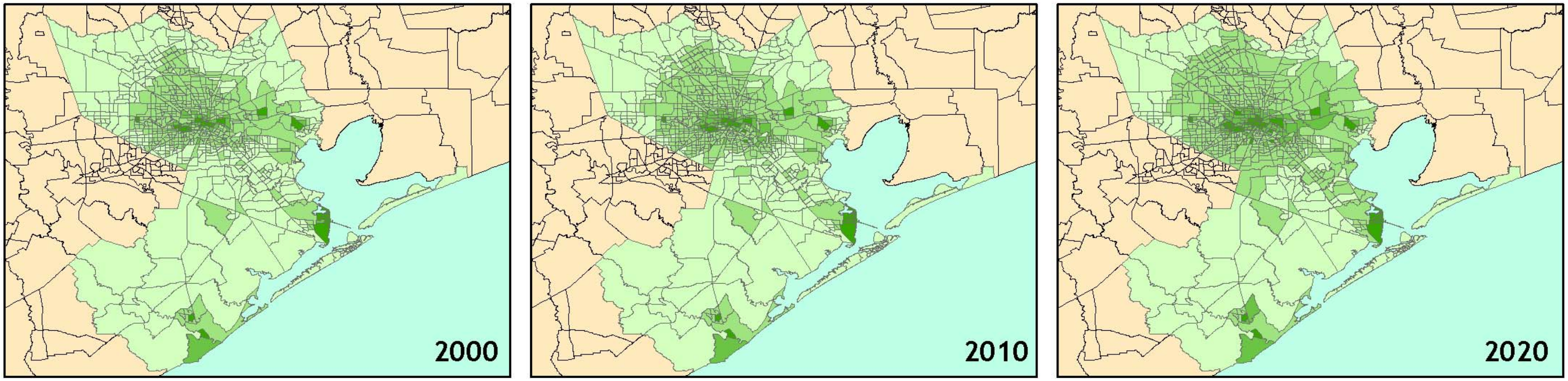
We applied valuation methods that are consistent the current economic literature and SAB advice concerning valuation of cancer-related outcomes. We valued fatal cancers using a value of statistical life (VSL) estimate, with an adjustment for medical costs associated with the period of cancer illness leading up to death. We valued non-fatal cancer cases using a per-case value based on SAB advice in a 2001 consultation on EPA’s arsenic in drinking water rule (USEPA, 2001a).

FIGURE ES-3: ESTIMATED CAAA-RELATED REDUCTIONS IN BENZENE CONCENTRATIONS IN THE HOUSTON METROPOLITAN AREA (*WITHOUT-CAAA* MINUS *WITH-CAAA*) - AERMOD AND HAPEM RESULTS

AERMOD RESULTS



HAPEM RESULTS



Reductions in Concentration 🟩 >2.5 µg/m³ 🟩 1.5 to 2.5 µg/m³ 🟩 0.5 to 1.5 µg/m³ 🟩 <0.5 µg/m³

Note: HAPEM results represent the estimated exposure concentration reduction for the median exposed individual in each census tract.

Tables ES-1 and ES-2 present our primary estimate for avoided fatal and non-fatal cases of leukemia due to CAAA-related changes in ambient benzene levels in the Houston area. Table ES-1 presents the number of expected annual cases avoided in each study year as well as the total cumulative avoided cases throughout the study period and the total cumulative avoided cases expected to occur after 2020, due to changes in benzene occurring within the study period. Table ES-2 shows the monetary value of the benefits of the avoided leukemia cases in the study period. Figure ES-4 illustrates the sensitivity of our results to alternative assumptions about the dose-response model.

TABLE ES-1: AVOIDED ANNUAL LEUKEMIA CASES (FATAL AND NON-FATAL) BY STUDY YEAR DUE TO CAAA-RELATED BENZENE EXPOSURE CHANGES IN THE HOUSTON AREA

STUDY YEAR	ANNUAL AVOIDED CASES OF LEUKEMIA		
	AVOIDED FATAL CASES	AVOIDED NON-FATAL CASES	TOTAL AVOIDED CASES
2000	0.03	0.02	**0.05**
2010	0.09	0.07	**0.2**
2020	0.2	0.1	**0.3**
Cumulative Cases Occurring Within the Study Period	**2**	**2**	**4**
Additional Cumulative Cases Occurring After 2020*	**1**	**1**	**2**
Total Cumulative Cases	**3**	**3**	**6**

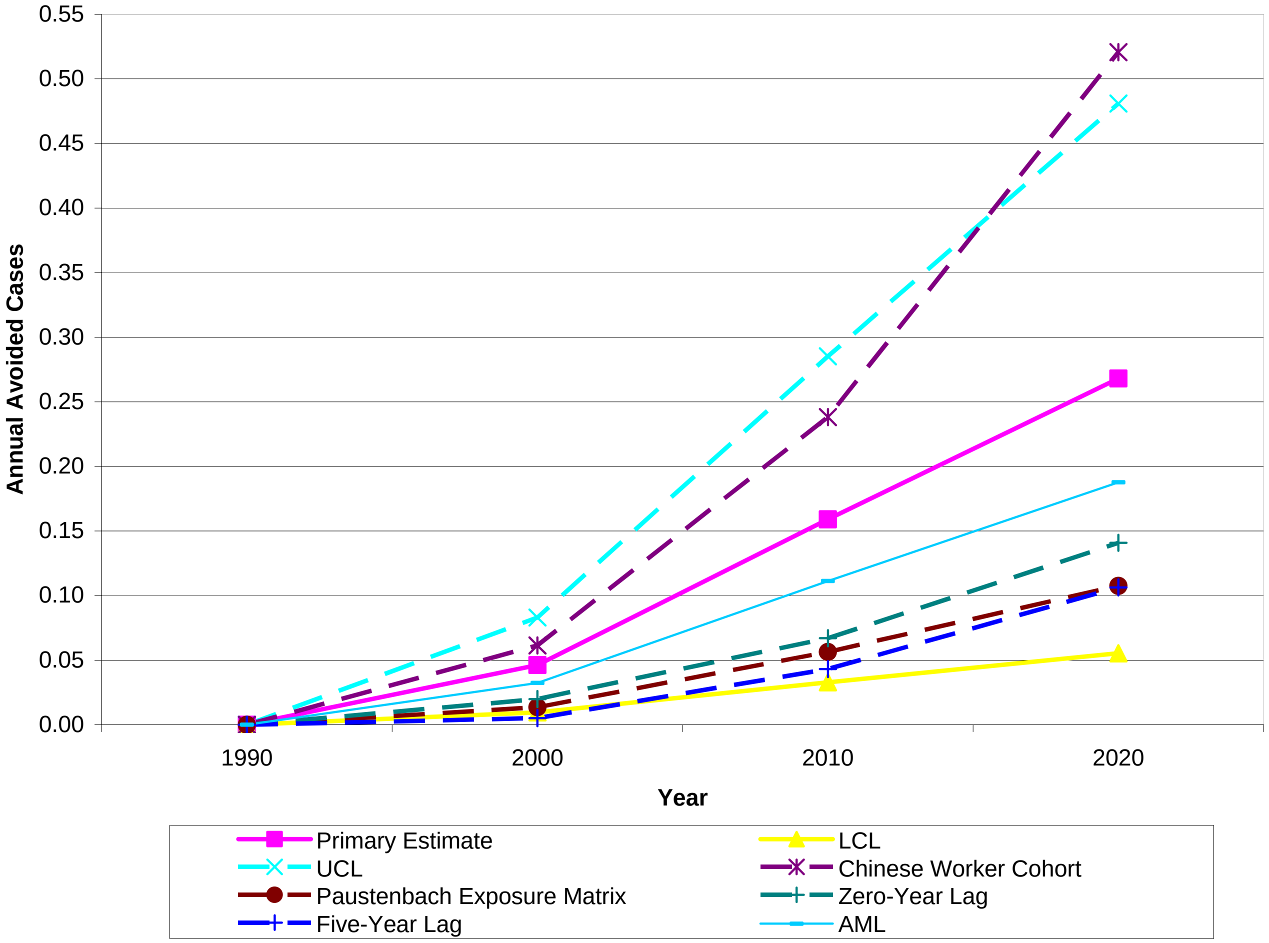
*Note: These avoided cases are due to changes in benzene exposure that occurred within the study period.

TABLE ES-2: TOTAL ANNUAL MONETARY BENEFITS THROUGH 2020 DUE TO CAAA-RELATED CHANGES IN BENZENE EXPOSURE IN THE HOUSTON AREA

STUDY YEAR	TOTAL BENEFITS (1990 NPV, MILLIONS OF 2006$, 5% DR)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
2000	$0.12	$0.01 – 0.06	**$0.13 - 0.18**
2010	$0.27	$0.01 – 0.13	**$0.28 - 0.40**
2020	$0.31	$0.01 – 0.15	**$0.32 - 0.46**
Cumulative Cases Occurring Within the Study Period	**$6.7**	**$0.32 - 3.3**	**$7.0 - 10**
Additional Cumulative Cases Occurring After 2020*	**$1.6**	**$0.08 - 0.8**	**$1.7 - 2.4**
Total Cumulative Cases	**$8.3**	**$0.40 - 4.1**	**$8.7 - 12**

*Note: These avoided cases are due to changes in benzene exposure that occurred within the study period, but occurred after 2020 due to lagging effects of these changes on leukemia risks.

FIGURE ES-4: ANNUAL AVOIDED CASES OF LEUKEMIA DUE TO CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA - PRIMARY ESTIMATE AND SENSITIVITY ANALYSES RESULTS



Note: We have linearly interpolated between the avoided leukemia estimates for each target year; however, the true shape of the curve between each of these points is uncertain.

In addition to the leukemia analysis, we evaluated the numbers of individuals likely to be exposed to benzene at levels exceeding EPA's chronic reference concentration (RfC) for benzene, which is based on changes in white blood cell counts, under the *With*-and *Without-CAAA* scenarios. We found no individuals exposed to benzene at concentrations exceeding the RfC in either the *With*- or *Without-CAAA* scenarios. We also conducted illustrative analyses of exposure and risk reductions to highly exposed subpopulations in the study area, and found potentially significant individual risk reductions due to the CAAA for individuals in these groups.

In summary, this case study demonstrates that the 1990 CAAA controls on benzene emissions are expected to result in reductions in the incidence of leukemia in the greater Houston area over the period 1990 to 2020. Key findings include:

- CAAA programs are expected to reduce benzene emissions across all source categories in the study area by thousands of tons per year, with the largest reductions in the point and non-point source category, followed by on-road and non-road sources;
- The largest reductions in benzene exposures are expected to occur in downtown Houston and the surrounding area, and in two areas with significant point sources: the Texas City area of Galveston County and southeastern Brazoria county;
- Reductions in benzene levels are expected to continue, and hence benefits are expected to increase in the latter decades of the study period, as engine and other capital stock turns over and the impact of CAAA controls on on-road and non-road mobile sources in the area increases;
- Primary benefit estimates indicate four fewer cases of leukemia would occur in the three-county area in the study period, two of which we expect would have been fatal. We also expect benefits from the benzene changes that occur between 1990 and 2020 will continue accruing through at least 2030, potentially avoiding another two leukemia cases between 2020 and 2030. We estimate the net present value (NPV) in 1990 of the two fatal and two non-fatal leukemia cases avoided is between $7 and 10 million in 2006 dollars, based on a five percent discount rate.
- 1990 CAAA controls on benzene are expected to significantly reduce individual leukemia risk levels for those living in census tracts with the highest estimated benzene levels by one to two orders of magnitude. For example, some risks in Brazoria County drop from an increased lifetime leukemia risk of 2 in ten thousand (i.e., 2×10^{-4}) to 3 in a million (3×10^{-6}). In four of the six census tracts with the highest risks, individual lifetime leukemia risks are reduced by at least 80 percent.
- Additional health benefits may accrue to individuals living in homes with attached garages. Back-of-the-envelope estimates of the benefits of CAAA-related benzene reductions in the garages of these homes suggest these benefits may be similar in magnitude to our primary estimate. Therefore, these results suggest that adding attached garage-related benefits to our primary estimate could result in an approximate doubling of our primary estimate.

To place these results in context, we note that this air toxics case study focuses only on a subset of the health effects associated with benzene exposure and does not include the total benzene emissions reductions achieved in the Houston area by the CAAA. As such, the case study does not provide a comprehensive assessment of current health effects resulting from benzene exposures in the Houston area; nor does it provide a full measure of the benefits that could be achieved by reducing current benzene emissions affecting the area. Additional caveats to consider when interpreting the results of this case study include:

- Recent studies in the Houston area suggest that emissions inventories such as the ones used in this case study may significantly underestimate local emissions of VOCs such as benzene from large point sources (e.g., refineries). To the extent that CAAA programs would reduce these emissions, we would not capture these benefits in the case study.
- The case study results include only overall leukemia effects associated with reductions in benzene emissions achieved by a subset of new controls implemented pursuant to the 1990 amendments to the Clean Air Act.
- Reductions from new programs established since we began this case study, especially the Mobile Source Air Toxics Rule, are not included in the analysis.
- Additional health effects that may be associated with benzene exposure but were not included in the quantitative results include other cancers, such as Hodgkin’s Lymphoma, and non-Hodgkin’s Lymphoma, multiple myeloma, and myelodysplastic syndrome; and potential non-cancer effects related to various hematological abnormalities, including aplastic anemia.
- Co-benefits of reducing air toxics, including reductions in ozone and particulate matter levels, are captured in the overall section 812 study but are not incorporated in the case study.

Despite the limitations of this case study, it successfully demonstrates a methodology that can serve as a useful tool in EPA's evolving HAP benefits assessment strategy. It can provide a comprehensive assessment of the impact of benzene controls from multiple CAAA Titles on cancer incidence in an urban population, using a combination of national and local data to conduct urban-scale modeling of air quality and health impacts. Further, the life-table model allows for more careful assessment of risk changes over time at the census tract level, incorporating local, age-specific baseline incidence data with age-specific exposure data and information on the lag between exposure changes and risk reductions.

Determining where this approach best fits within EPA's HAP benefits assessment strategy will require additional analysis and evaluation to determine the added value of the detailed, urban-scale approach, as well as the potential pool of HAPs suitable for assessment via the damage-function approach for cancer and/or non-cancer effects.

CHAPTER 1 | INTRODUCTION

Section 812 of the Clean Air Act Amendments of 1990 (CAAA) requires the U.S. Environmental Protection Agency (EPA) to perform periodic, comprehensive analyses of the total costs and total benefits of programs implemented pursuant to the Clean Air Act (CAA). The first analysis required was a retrospective analysis, addressing the original CAA and covering the period 1970 to 1990. The retrospective was completed in 1997. Section 812 also requires prospective cost-benefit analyses, the first of which was completed in 1999. The prospective analyses address the incremental costs and benefits of the CAAA. The first prospective analysis covered implementation of the CAAA over the period 1990 to 2010.

EPA's Office of Air and Radiation (OAR) began work on the second prospective study in 2003 with the drafting of an analytical plan for the study. One of the objectives of the analytical plan was to address past comments from EPA's Science Advisory Board Advisory Council for Clean Air Compliance Analysis (SAB Council) concerning treatment of hazardous air pollutants (HAPs) in the previous 812 studies. Assessing the benefits of Clean Air Act controls on the 188 HAPs listed in Title III of the CAAA is much more challenging than analyzing the benefits associated with criteria pollutant reductions, which are the focus of the main 812 benefit/cost analysis. Difficulties include gaps in the toxicological database; difficulty in designing population-based epidemiological studies with sufficient power to detect health effects; limited ambient and personal exposure monitoring data; limited data to estimate exposures in some critical microenvironments; and insufficient economic research to support valuation of the types of health impacts often associated with exposure to individual HAPs. As a result, EPA's efforts to characterize the benefits of HAP reductions in prior 812 analyses have been only partially successful. The SAB Council criticized an analysis of National Emissions Standards for Hazardous Air Pollutants (NESHAP) regulations conducted for the retrospective analysis as substantially overstating benefits, with particular note made of the use of "upper bound" dose-response relationships (i.e., the cancer potency factor used for standard setting).

EPA made a second attempt to incorporate air toxics benefits, in the first prospective analysis (USEPA, 1999a), but the SAB Council found that the national air quality and exposure model proposed would not yield estimates suitable for benefits analysis. In 2001, the SAB Council proposed that EPA undertake a HAP benefits case study, and suggested benzene as a candidate pollutant. The SAB recommended benzene in part because of the wealth of available national ambient concentration data from monitors. The SAB believed that an 812 analysis using the available benzene data would:

- Identify limitations and gaps in the database of air toxics health impact functions;
- Provide an estimate of the uncertainties in the analyses and perhaps provide a reasonable lower bound on potential health benefits from control; and
- Provide a scientific basis for deciding whether there is merit in pursuing a greater ability to assess the benefits of air toxics (USEPA, 2001b).

In response to these comments, EPA conducted a metropolitan scale case study of the benefits of CAAA controls on benzene emissions to accompany the main 812 analysis. This report describes the methodology and results of that analysis.

1.1. PURPOSE AND SCOPE

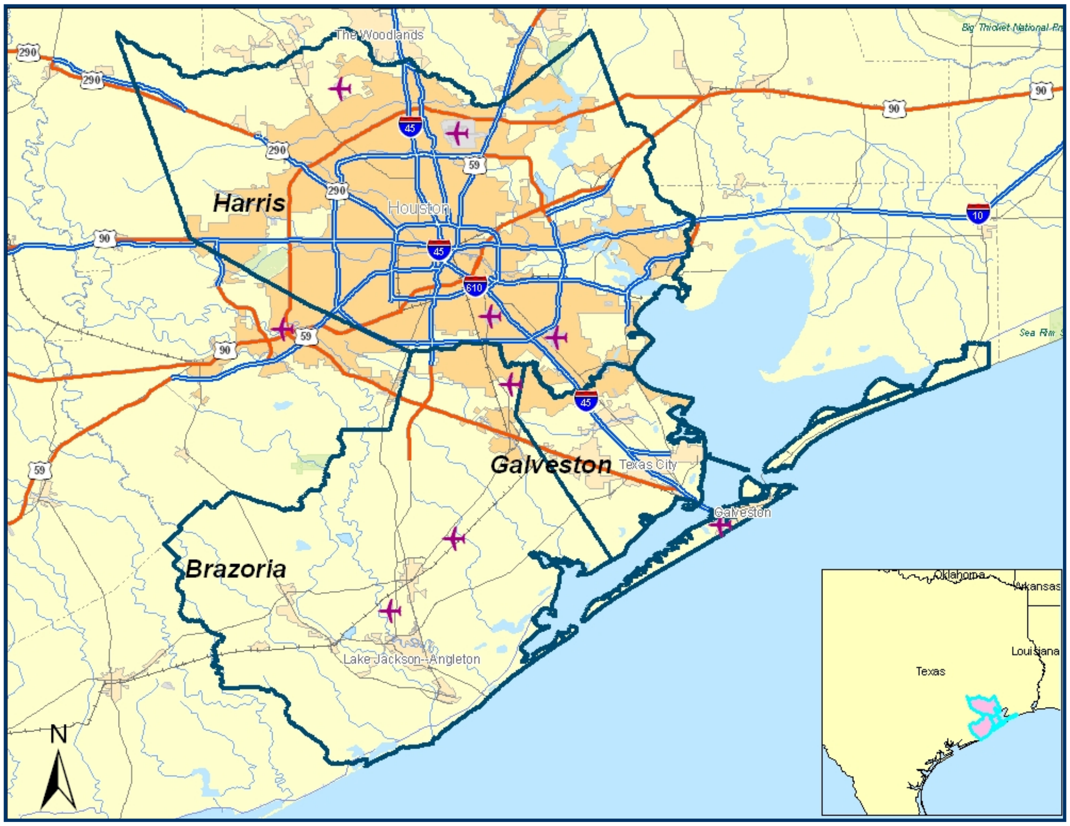
This case study has two main objectives. The first is to demonstrate a methodology that generates human health benefits resulting from CAAA controls on a single HAP in an urban setting, while highlighting key limitations and uncertainties. The second is to provide a basis for considering more broadly the value of such an exercise for HAP benefits characterization nationwide. This case study is not intended to provide a comprehensive assessment of the benefits of benzene reductions due to the Clean Air Act.

We selected the Houston-Galveston area for the case study, a metropolitan area with a large population (a total of 3.4 million in 2000, with nearly 3 million people in Harris County alone) and significant benzene emissions from both on-road mobile sources and large industrial point sources such as petroleum refineries.

Figure 1 displays the study area for this analysis. The study area encompasses three counties – Harris, Galveston, and Brazoria. The selection of these counties represents a balance of resource concerns with comprehensiveness – according to EPA's 1999 National Emissions Inventory (NEI), these three counties contribute 99% of the point source emissions in Houston metropolitan area. The study area also captures significant major roadways, including Interstate 45 and the loops surrounding downtown Houston, Houston's major airports (Bush/Intercontinental and Hobby International), the Port of Houston and the Houston Ship Channel, the refineries of Texas City, and major chemical manufacturing and petroleum refining facilities located in southeastern Brazoria county.

The timeframe for this analysis, 1990 through 2020, matches that used in the criteria pollutant analysis of the second prospective Section 812 study. In addition to the base year, 1990, we model results for three target years, 2000, 2010, and 2020. For each of the target years, we model benzene exposures and health impacts under two scenarios, one reflecting the impacts of regulatory programs enacted in response to the 1990 CAAA, and one assuming no additional benzene pollution control activity beyond the regulatory requirements existing in 1990.

FIGURE 1: BENZENE CASE STUDY AREA



1.2 ORGANIZATION OF THIS DOCUMENT

The remainder of this document is divided into three chapters. Chapter 2 describes our analytical approach to the benefits analysis. Chapter 3 presents the results for the various steps in the analytical chain (i.e., emissions, air quality, exposure concentrations, and health benefits). Chapter 4 presents our conclusions and a discussion of the usefulness of the methods described in this report for broader HAP benefits analysis. The report also includes five appendices. Appendix A provides a detailed description of the emissions modeling conducted by E.H. Pechan and Associates (Pechan). Appendix B describes in detail the air quality modeling performed by EPA. Appendix C presents a health effects literature review memo conducted in support of the health benefits modeling approach, Appendix D presents additional details about the health benefits model, and Appendix E provides the equations used in an analysis of attached garage benzene exposures.

CHAPTER 2 | ANALYTICAL APPROACH

This chapter describes the approach we applied to estimate the health benefits of reduced benzene emissions in Houston between 1990 and 2020 resulting from the Clean Air Act Amendments of 1990 (CAAA). We conducted this benefits analysis using the standard approach applied in the main 812 criteria pollutant analysis, which includes the following five steps:

1. Scenario Development
2. Emissions Estimation
3. Air Quality and Exposure Modeling
4. Health Effects Modeling
5. Valuation

We provide in the following sections a description of our methods for each step in the analytical chain. More detailed information for steps 2 through 4 may be found in Appendices A through D of this document.

2.1 SCENARIO DEVELOPMENT

To assess benefits of CAAA-related benzene reductions, we adopted an approach consistent with the main criteria pollutant analysis in the 812 study. Our benefit analysis is based on construction and comparison of two distinct scenarios: a *Without-CAAA* scenario and a *With-CAAA* scenario. This case study then estimated the differences between the health outcomes associated with these two scenarios.

The *Without-CAAA* scenario essentially freezes federal, state, and local air pollution controls at the levels of stringency and effectiveness that existed in 1990. This scenario is consistent with the baseline for the main 812 analysis. The *With-CAAA* scenario assumes that all federal, state, and local rules promulgated pursuant to, or in support of, the 1990 CAAA were implemented. This scenario includes all current and currently anticipated regulations that affect benzene emissions resulting from the amended clean air act issued in 1990. It includes the regulations listed in Table 1. We note that the Mobile Source Air Toxics (MSAT) rule issued by EPA on February 26, 2007, was not included in this scenario, because the rulemaking was still in progress at the time of this analysis. We expect the MSAT rule to further reduce benzene emissions under the *With-CAAA* scenario beyond what we report in this case study.[2]

---

[2] Other planned control programs on small spark ignition engines, including locomotive and marine engines, may also reduce benzene further (http://www.epa.gov/otaq/regs/nonroad/marinesi-equipld/420f07032.htm).

TABLE 1: BENZENE CASE STUDY *WITH-CAAA* SCENARIO SUMMARY, BY TITLE

Title I	Any effects of Title I will be expressed through state implementation plan (SIP) requirements, such as (enhanced) I/M programs, transportation control measures, and other VOC controls. These requirements are dependent on the ozone non-attainment status of the case study area(s).
Title II	**<u>Tailpipe standards</u>** **On-road** Tier 1 Standards (phased in 1994 to 1997) National Low-Emission Vehicle (NLEV) program – voluntary bridge between Tier 1 and Tier 2 Tier 2 Standards take effect in 2004 Heavy Duty Engine/Diesel Fuel Rule - New emission standards – 2007 model year, new fuel standards 2006 **Non-road** Federal Phase I and II compression ignition (CI) engine standards Federal Phase I and II spark ignition (SI) engine standards Federal locomotive standards Federal commercial marine vessel standards Federal recreational marine vessel standards **<u>Evaporative Emissions</u>** Stage II Vapor Recovery Systems (Section 182) Onboard Refueling Vapor Recovery (Section 202; 1998 model year and on) Evaporative Test Procedure **<u>Fuel Regulations</u>** Reformulated Gasoline (RFG) Standards (1995 on) Phase II – (2000 – present) – benzene requirements essentially unchanged Summertime Volatility Requirements for Gasoline (Phase II – 1992 on) Fuel Sulfur Limits *2007 Mobile Source Air Toxics (MSAT) Rule NOT included.
Title III	**<u>Maximum Achievable Control Technology (MACT) Standards</u>** The *With-CAAA* scenario included MACT standards that would be expected to have a significant effect on future-year benzene emissions in the Houston area. These standards include: Oil and Natural Gas Production: 7-Year MACT Petroleum Refineries: 4-Year MACT Gasoline Distribution: 4-Year MACT Pulp and Paper Production: 7-year MACT Municipal Landfills: 10-year MACT Natural Gas Transmission and Storage: 10-year MACT Publicly Owned Treatment Works (POTW) Emissions: 7-year MACT Coke Ovens: Pushing, Quenching, & Battery Stacks: 4-year MACT Synthetic Organic Chemical Manufacturing Industry Hazardous Organic NESHAP (SOCMI HON): 2-year MACT

This approach requires two simplifying assumptions. First, we assumed, as noted above, that under the *Without-CAAA* scenario regulations are frozen at 1990 levels, and that no additional voluntary, state, or local pollution control activities occur in the Houston area beyond regulations "on the books" as of 1990. Second, we assumed that the distribution of population and economic activity is the same under both scenarios. Thus, for example, as air quality worsens under the *Without-CAAA* scenario, we did not attempt to model the movement of individuals out of the study area. While migration might in fact occur in response to a degradation in air quality, estimation of the extent of this migration would have required speculation that would not necessarily have been superior to our chosen approach.

Although this is a HAP case study, we did not analyze benefits specific only to Title III of the CAAA (the Title that specifically focuses on HAPs), because doing so would have ignored significant benefits related to reductions of benzene emissions from mobile and stationary sources. Instead, the difference between the *With*- and *Without-CAAA* scenarios for benzene in Houston reflects all CAAA regulations that affect benzene emissions.

2.2 EMISSIONS

We estimated benzene emissions in the Houston-Galveston study area for four source categories: point, non-point, on-road, and non-road. For each category, we generated emissions estimates for the 1990 base year and for three target years (2000, 2010, and 2020) under both the *With-CAAA* and *Without-CAAA* scenarios.

Our overall approach involves estimating emissions by backcasting or forecasting historical emissions data based on the expected growth in emissions-generating activities over time, adjusted for the impact of future year control assumptions under each scenario. For the *With-CAAA* scenario we estimate emissions for the three target years by adjusting benzene emissions from EPA's 2002 National Emissions Inventory (NEI). For the *Without-CAAA* scenario, we generate projections for the three target years by adjusting the 1990 benzene emissions from EPA's 1990 NEI.

The emissions estimates thus depend primarily on the following three elements: 1) the data and methods used to establish the historical year inventories; 2) the indicators used to forecast growth in emissions activities and emission control effectiveness; and 3) the specific regulatory programs modeled under the *With*- and *Without-CAAA* scenarios. We have included a detailed report (E.H. Pechan and Associates, 2006) describing the approach to each of these elements for each source category in Appendix A. We provide an overview of the methods used for each sector below.

2.2.1 POINT SOURCES

We estimated point source emissions in 1990 in the study area using EPA's 1990 National Emissions Inventory (NEI) for HAPs, which was recently revised by EPA.[3] This inventory also served as the base year file for estimating *Without-CAAA* scenario emissions for 2000, 2010, and 2020. We estimated point source benzene emissions for the study area for the 2000 *With-CAAA* scenario by backcasting from the EPA 2002 NEI, and this served as the base year emissions file for preparing 2010 and 2020 *With-CAAA* scenario emission estimates.[4]

For the 1990 NEI, EPA established a hierarchy of preferred data sources in order to prepare the 1990 NEI for HAPs, listed below in order of preference:

- Maximum Achievable Control Technology (MACT) data from EPA's Emission Standards Division (ESD);
- Data developed by state and local air agencies;
- Data from inventories developed by EPA's Emission Inventory Group to support requirements of Sections 112(c)(6) and 112(k); and
- Emissions reported in the Toxics Release Inventory (TRI), and emissions that the Emission Inventory Group generated using emission factors and activity factors.

Nearly 90 percent of the 1990 point source emissions data for the study area came from the first two categories.

The sources of data in the NEI for benzene in the Houston-Galveston area are as follows:

- State data,
- Industry data;
- Data gathered by EPA's Emission Standards Division while developing MACT standards;
- TRI data; and

---

[3] The original baseline 1990 NTI was a county-level inventory for all source categories. The newly released 1990 NEI for HAPs was created by converting the county-level emission estimates to facility-specific estimates for as many sources as possible. Locational data and stack parameters were added, and additional estimates were developed for missing MACT source categories and HAPs.

[4] We also considered basing the 2000 *With-CAAA* scenario benzene emissions estimates on Texas Council on Environmental Quality (TCEQ) emissions data generated for use in the Texas Air Quality Study (AQS) 2000 study. EPA's Science Advisory Board encouraged EPA to investigate these data as an alternative to the 1999 NEI, which we had originally proposed to use. We chose not to use the Texas AQS data for several reasons. First, it would have required significant adjustments to generate year round emissions estimates, as it only provided data for an August-September 2000 Houston area modeling episode. Second, it lacked the control device information that assists in making emission forecasts to future years. Furthermore, subsequent to the SAB consultation, the 2002 NEI was issued, and the TCEQ submittal for the 2002 NEI reflected improved point source emissions estimates over the 1999 submission. The availability of this improved inventory, combined with the limitation of the Texas AQS data led us to select the 2002 NEI.

- Electricity generating unit (EGU) data developed from information by the U.S. Department of Energy (DOE) and EPA's Clean Air Markets Division (CAMD).

State data comprised over 87 percent of the 2002 point source emissions data, with the rest primarily coming from TRI.

2.2.2 EMISSIONS ACTIVITY FACTORS

When estimating point source category emissions for future years, we applied emission activity factors that reflect the projected ratios of 2000, 2010, and 2020 emission activity to 1990 emission activity (for *Without-CAAA* case emissions estimation) and the ratios of 2000, 2010, and 2020 emission activity to 2002 emission activity (for *With-CAAA* case emissions backcasting and forecasting).[5] We developed emission activity levels for energy-producing and energy-consuming source categories from historical/forecast energy production/consumption data. Because it is not feasible to develop estimates of actual emission activity levels for every non-energy related source category, we derived growth factors for these source categories from surrogate socioeconomic indicator data that are more readily available than emission activity data.

In keeping with past EPA practice, we relied on energy data from DOE's Energy Information Administration (EIA) to backcast and forecast energy consumption and energy production emission source categories. To reflect the 1990 to 2000 trend in energy consumption for source categories, Pechan generally relied on historical time-series energy data for Texas from an EIA energy consumption database (EIA, 2005a). For Crude Oil and Natural Gas Production source categories, Pechan obtained 1990 and 2000 Texas relevant activity data from another EIA source that provided the number of operating oil well days (used for Crude Oil Production) and the number of operating gas well days (used for Natural Gas Production) (EIA, 2005b). For source categories that describe railroad and marine distillate fuel consumption emission processes, Pechan obtained 1990 and 2000 consumption estimates for Texas from an EIA distillate fuel data resource (EIA, 2005c).

When projecting activity for future years, we relied on projections of U.S. energy supply, demand, and prices through 2025, which are published by EIA in its *Annual Energy Outlook 2005 (AEO 2005)* (EIA, 2005d). We applied *AEO 2005* West South Central (WSC) region forecasts to project changes in Houston-Galveston area emissions activity (Texas is included in the WSC region). For example, Stage II (Gasoline Vehicle Refueling) emission activity is projected using *AEO 2005* projections of WSC region transportation sector motor gasoline consumption. We relied on national energy forecasts when *AEO 2005* only provided national projections for the energy growth indicator of interest.

---

[5] We also applied this approach for projecting non-road source categories that are not incorporated into EPA's NONROAD emissions model.

Because population growth and the performance of the U.S. economy are two of the main determinants of energy demand, the EIA also prepares socioeconomic projections. This study relies on *AEO 2005* historical and forecast socioeconomic data as surrogates for emission activity level changes for most non-energy source categories.[6]

2.2.3 NON-POINT SOURCES

Non-point (formerly "area source") emissions were projected for both the *With-CAAA* scenario (2010 and 2020) and the *Without-CAAA* scenario (2000, 2010, and 2020). The draft 2002 NEI was used as the initial base for the *With-CAAA* scenario, while the 1990 NEI for HAPs inventory was used as the initial base for the *Without-CAAA* scenario.

We conducted a ranking of non-point and non-road benzene emitting categories (i.e., Source Classification Codes (SCCs)) for the 3-county Houston-Galveston area based on benzene emissions reported for EPA's draft 2002 NEI. Based on this ranking, we identified five priority SCCs on which to focus this analysis: gasoline marketing, commercial marine loading, bulk terminals, pipeline facilities, and commercial marine diesel engines.[7] Adjustments to these emissions data to generate projections for future years in each of the two scenarios followed the procedure described in the point source section. Detailed descriptions of additional adjustments to NEI data for use in this analysis are provided in Appendix A.

2.2.4 ON-ROAD SOURCES

For the on-road source category, we calculated emissions estimates that are linked to specific roadway segments (i.e., link-level estimates) for the 1990 base year and all three target years under the *With-* and *Without-CAAA* scenarios. Link-level values have been applied in other urban-scale air quality analyses (e.g., EPA, 2002, Stein et al., 2002, Hao et al, 2002) and can provide improved emissions resolution over grid-based methods for air quality modeling at the census block group level. We generated hourly link-level emissions data by season for each year/scenario combination.

The inputs to the on-road emissions estimation process include estimates of vehicle miles traveled (VMT) and emissions factors (e.g., grams of benzene per mile traveled) for specific vehicle types and driving conditions. We prepared link-level VMT data from VMT data files prepared in 2005 for the Houston area by the Houston-Galveston Area Council and further processed by the Texas Transportation Institute (TTI). We obtained VMT data sets for 2002, 2009, and 2012, and adjusted the data as necessary to match the temporal needs of the Section 812 study. For example, we adjusted the VMT data,

---

[6] For four source categories – Residential Wood Fireplaces and Wood Stoves, Aircraft, Forest Wildfires, and Prescribed Burning for Forest Management - we opted to use non-*AEO 2005* surrogates for projecting emissions activity. We applied methods to derive growth factors for these categories that are consistent with those used in past EPA analyses, such as the Clean Air Interstate Rule Regulatory Impact Analysis (RIA). The specific approaches we applied are described in Appendix A.

[7] Portable fuel containers are another significant source of non-point/non-road benzene emissions that were not included in this assessment.

originally developed for an August/September ozone modeling episode, from the specific modeling period to the four seasons, using adjustment factors provided by TTI. We then allocated the VMT by vehicle type and adjusted the data to the study years for this analysis. We conducted the study year adjustment by calculating the average annual VMT growth rate between the two years of VMT data nearest to the year of interest (e.g., 2002 and 2012 for 2020) and then applying that rate to interpolating or extrapolating VMT for that year.

Emission factors were calculated using EPA's MOBILE6.2 model.[8] Where possible, local input data for the Houston area, as provided by the Texas Commission on Environmental Quality (TCEQ), were used in the development of the MOBILE6.2 input files. Emissions factor inputs include registration distributions of vehicles by age, diesel sales fractions, inspection and maintenance (I/M) and Anti-Tampering Program (ATP) inputs, temperature, and fuel characteristics and properties. Details on the specific data used for each of these input categories can be found in Appendix A. Once the input files were prepared, we ran MOBILE 6.2 for the 1900 base year and for 2000, 2010, and 2020 under the *With*- and *Without*-*CAAA* scenarios. For the *With-CAAA* scenarios, MOBILE6.2 generated emissions factors that reflect the impact of I/M programs and ATPs instituted in the study area after 1990, as well as emissions requirements and fuel programs in place in the year being modeled. For the *Without*-*CAAA* scenarios, we ran MOBILE 6.2 using 1990 fuel characteristics and the "NO CAAA" command, which excluded the effects of national CAAA programs on emission factors.

2.2.5 NON-ROAD SOURCES

To develop non-road benzene emission estimates from in the Houston, Texas area, we first used EPA's NONROAD2004 model to generate volatile organic compounds (VOC) exhaust and evaporative VOC emissions output from non-road sources for the 1990 base year and future years under both the *With*- and *Without*-*CAAA* scenarios (USEPA, 2004a).[9] We obtained VOC emissions estimates for the following model equipment categories: recreational vehicles, farm and construction machinery, lawn and garden equipment, aircraft and rail support equipment, and other industrial and commercial applications. Aircraft, commercial marine and locomotive emissions, which are not modeled by NONROAD, were included in the non-point area source portion of the emissions inventory.

---

[8] Analysis for the recent MSAT rule found that cold start emissions for Tier 1 and later vehicles are much larger than estimated by MOBILE6, suggesting a potential downward bias on emission reduction estimates for this category; however, the impact of these emissions in Houston is likely smaller than in colder climates.

[9] The NONROAD2004 model was released by EPA's Office of Transportation and Quality (OTAQ) in May 2004. This version of the model incorporates all Federal engine standards, with the exception of the large spark-ignition evaporative standards. VOC reductions from this standard were applied outside of the NON-ROAD model, as described in Appendix A. A recent revision to NONROAD (NONROAD2005) includes new evaporative emission categories, such as tank and hose permeation, and revised hot soak emission estimates, which increase the inventory. These revisions are not included in our analysis.

To estimate the specific benzene emissions associated with NONROAD's various categories of VOC emissions, we compiled engine-specific benzene speciation factors for exhaust and evaporative emissions from EPA's National Mobile Inventory Model (NMIM) (USEPA, 2005a).[10] We then multiplied the SCC-level VOC emissions estimates by these factors to produce estimates of benzene emissions from non-road sources.[11] The specific benzene speciation factors applied can be found in Appendix A.

We employed a revised NONROAD model growth file with region-specific growth rates, consistent with the main criteria pollutant analysis of the Section 812 Prospective study.[12] Input files were prepared for Brazoria, Galveston, and Harris counties to reflect the appropriate temperature and fuel inputs for the *With-CAAA* scenario runs.[13] In addition, fleet emission rate inputs were modified to remove the effect of CAAA-related standards for the *Without-CAAA* runs. Using county-specific input files, NONROAD model runs were performed to generate seasonal emission estimates for each scenario year. Seasonal emissions were then summed to estimate annual emissions at the county and SCC level for each scenario/year.

2.3 AIR QUALITY AND EXPOSURE MODELING

The air quality modeling (AQM) step links emissions changes within the three-county study area to changes in atmospheric concentrations of benzene. It replicates dispersion and transport of emitted benzene through the atmosphere to generate a set of estimated ambient benzene concentrations at the census tract level. When combined with information about the time-activity patterns of an exposed population, the ambient AQM estimates can be converted to estimates of individual exposure concentrations for that population.

We applied EPA's American Meteorological Society/U.S. EPA Regulatory Model (AERMOD) dispersion modeling system (U.S. EPA 2004b) to convert emissions estimates to ambient benzene concentrations in the Houston-Galveston study area in the base and target years under the *With-* and *Without-CAAA* scenarios.[14] AERMOD is a state-of-the-art steady-state Gaussian plume model that is one of EPA's preferred models

---

[10] Evaporative hydrocarbon emissions as calculated by NONROAD are comprised of crankcase, diurnal, spillage, and vapor displacement components.

[11] No benzene emission factors were available (or applied) for Liquified Petroleum Gas or Compressed Natural Gas-fired equipment.

[12] The procedures used to develop the regional growth rates are described in the Section 812 Prospective report (Pechan, 2005a).

[13] Input parameters for Brazoria, Galveston, and Harris counties were developed that reflected local and national fuel programs for the *With-CAAA* scenario runs for 2000, 2010, and 2020. Local inputs, including seasonal reduced vapor pressure (RVP) limits, oxygenated fuel specifications for reformulated gasoline, and Stage II programs were available from EPA's NMIM county database (USEPA, 2005a). Federal gasoline and diesel fuel sulfur levels were incorporated as well.

[14] We also considered using the Community Multiscale Air Quality (CMAQ) model to estimate ambient benzene concentrations. However, benzene is a relatively stable compound and therefore the ability of CMAQ to account for photochemical processes was not necessary. In addition, the AERMOD model is able to provide finer spatial scale resolution.

for regulatory analyses of this scale; it handles multiple sources, incorporates building downwash, has flexibility in receptor location choices, and also includes the option to vary emissions by season and hour of day. We fed the AERMOD output into EPA's Hazardous Air Pollutant Exposure Model, Version 6 (HAPEM6) to generate benzene exposure concentrations.

The next three sections describe the AERMOD modeling approach, and the fourth covers the HAPEM exposure modeling.

2.3.1 AQM MODEL INPUTS

Inputs to the model included a receptor grid (i.e., the geographical locations at which concentrations are to be estimated); the emissions data from the previous step, which were processed to conform with AERMOD requirements; meteorological data; land use and elevation data; and information on background levels. Detailed information about the development of each input can be found in Appendix B (note that Appendix B uses the term "area and other" to refer to non-point source emissions). We present below a brief overview of how each of these inputs was handled:

- **Modeling Domain/Receptors.** The modeling domain matched the three-county study area; we located receptors at census block group centroids.[15] We also placed some receptors at benzene monitoring locations for the purpose of model evaluation.
- **Emissions Data.** We employed EPA's Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP, Version 3.0, USEPA, 2004d) to process the seven emissions inventories developed by Pechan (2006) into the emissions input files required by AERMOD. The emissions processing required two steps. First, some of the emissions data required additional modifications prior to input into EMS-HAP, such as development of some source characteristics needed by AERMOD. Details of the emissions pre-processing can be found in Chapter 3 of Appendix B. Once the pre-processing was complete, we ran the emissions profiles through EMS-HAP to generate spatially and temporally allocated emissions input files appropriate for use with AERMOD. Additional information about EMS-HAP processing can be found in Chapter 4 of Appendix B.
- **Meteorological Data**. We prepared meteorological data for two years, 1990 and 2000. We input the 1990 meteorological data for the 1990 AERMOD simulation. We input the year 2000 data for all the other simulations. We used the AERMOD Meteorological Preprocessor (AERMET) (U.S. EPA, 2004c) to process the National Weather Service (NWS) data for both 1990 and 2000.

---

[15] A census block is a subdivision of a census tract. It is the smallest geographic unit for which the Census Bureau tabulates sample data. A block group consists of all the blocks within a census tract with the same beginning number (U.S. Census Bureau, 2007). For the 1990 simulation, the receptors were the 1990 census block group centroids, giving a total of 2,429 receptors. For all other AERMOD simulations in the study (2000, 2010, and 2020), the 2000 census block group centroids were chosen as the receptors, for a total of 2,285 receptors.

- **Land Use and Elevation Data.** We used data on land use to designate sources as urban or rural for dispersion modeling purposes. The urban/rural designation is important for AERMOD modeling when assigning deposition parameters. For non-point and non-road sources, excluding airport emissions, we assigned sources the land use designation of the census tracts to which the emissions were assigned during spatial allocation in EMS-HAP. We assigned each point source the land use designation of the closest tract. We modeled link-level on-road emissions as rural sources. This is consistent with previous studies in Houston (U.S. EPA, 2002a). We also modeled non-point and non-road airport related emissions as rural sources. Because the terrain is relatively flat over the Houston area, we ran the AERMOD simulations using the flat terrain option (i.e., we assumed sources and receptors are at the same elevation).
- **Background.** We added background concentrations to AERMOD modeled concentrations at each receptor (block group centroids) in a post-processing step to account for benzene contributions from sources outside the study area. We assigned background concentrations of benzene for all years and modeling scenarios based on the 1999 county specific background concentrations as used for the 1999 National Air Toxics Assessment (NATA, USEPA 2001b).[16] We applied the same background concentration to every receptor in a given county.

2.3.2 AQM MODEL RUNS

We performed seven model runs using AERMOD Version 04300 (one for the 1990 base year and two for each target year – one under the *With-CAAA* scenario and one under the *Without-CAAA* scenario). (The control options used for each run can be found in Appendix B, Table 12.).[17] For each model run, we generated hourly, daily, and annual average concentration output files for each source category (major, non-point, on-road, non-road, and total).[18] The hourly concentrations from AERMOD were then input into the HAPEM6 model (described in the next section), to generate exposure concentrations that reflect the influence of the activity patterns of the exposed population.

2.3.3 AQM MODEL EVALUATION

We performed an evaluation of the AERMOD results by comparing modeled concentrations to observed concentrations. In addition to the census block group centroids, we estimated daily and annual average model concentrations at monitor locations. We performed model-to-monitor comparisons for the year 2000 AERMOD

---

[16] For details about the 1999 background values see http://www.epa.gov/ttn/atw/nata1999/background.html or Battelle (2003).

[17] Receptors were the census block group centroids (the 1990 census block group centroids for the 1990 and the 2000 census block group centroids for all other years).

[18] Appendix B refers to non-point emissions as "area and other."

results using monitor observations obtained from EPA's Air Toxics Archive.[19] We identified 15 monitor locations available for comparison, mostly in southern Harris County (See Appendix B). We were unable to conduct a comparison for 1990, because only one benzene monitor existed in the study area at that time.

2.3.4 BENZENE EXPOSURE CONCENTRATION MODELING

Following completion of the AERMOD runs, we estimated time-weighted average benzene exposure concentrations for the study populations using the Hazardous Air Pollutant Exposure Model, Version 6 (HAPEM6) and the hourly ambient benzene concentration output from AERMOD. HAPEM assesses average long-term inhalation exposures of the general population, or a specific sub-population, over spatial scales ranging from urban to national. HAPEM6 tracks representatives of specified demographic groups as they move among indoor and outdoor MEs and among geographic locations.[20] The estimated pollutant concentrations in each ME visited are combined into a time-weighted average concentration, which is assigned to members of the demographic group (ICF, 2007). The model uses four main sources of information to calculate exposure: population data from the 2000 US Census; population activity data from the Consolidated Human Activity Database (CHAD) (Glen et al., 1997); commuting data from the 2000 Census; air quality data from AERMOD; and data on concentrations levels in MEs versus ambient levels. As part of the ME evaluation, algorithms accounting for the gradient in concentrations of primary mobile source air toxics within 200 meters of major roadways are used, which is an addition since the previous version of HAPEM (Version 5).

The HAPEM6 output from the runs performed for this study consisted of average annual exposure for an individual at the census tract level in each of six demographic groups. The demographic groups were determined by age (0-1; 2-4; 5-15; 16-17; 18-64; and ≥65 years). Contributions to ambient concentrations were calculated for the following source sectors: point ("major" in Appendix B), non-point ("area and other" in Appendix B), on-road, non-road, and background (USEPA, 2007a). Concentrations were provided for the 1st, 5th, 10th, 25th, 75th, 90th, 95th, and 99th percentiles, average, and median concentration for each source category, age group, and census tract in each of the target years for this study (1990, 2000, 2010, 2020).

2.4 HUMAN HEALTH EFFECTS ESTIMATION

This section presents our approach for estimating avoided adverse health effects in humans resulting from reductions in exposures to benzene in ambient air and in various MEs in the Houston area. We first review the epidemiological evidence evaluating potential health effects of benzene exposure and present the health endpoints included in

---

[19] EPA's Air Toxics Archive (http://vista.cira.colostate.edu/atda) contains multiple years of monitor observations for multiple HAPs across the U.S. The Archive contains a program that performs quality assurance on daily monitor observations and calculates an annual average concentration for each valid monitor.

[20] The model includes a total of 14 MEs, such as residential, school, office, public transit, and service station.

the human heath effects estimation. Based on the available evidence, we have focused our evaluation on the epidemiological evidence examining the link between benzene and leukemia. We next describe our selection of dose-response model for our analysis and review the exposure modeling conducted for the study population. We then describe our leukemia risk model, which employs a life-table approach to risk analysis. We close the section by describing our approach for estimating non-cancer health effects and describe ancillary illustrative analyses of high-exposure subpopulations, including residents living in high exposure census tracts, residents living near roadways, and residents with attached garages.

2.4.1 SELECTION OF HEALTH ENDPOINTS

Benzene is a very well studied chemical with a substantial database of epidemiological data associating it with leukemia. There is also limited evidence supporting a link between benzene and other health effects, such as other cancers and non-cancerous effects. IEc conducted an extensive literature review of the heath effects of benzene exposure to identify health endpoints for which the benefits of benzene reductions could be estimated. Note that this literature review was completed in early 2005. Therefore, our results do not reflect the findings of additional studies completed since that date. This section describes the health endpoints selected for the human health effects analysis as a result of that review and our rationale for including them. Additional details may be found in Appendix C.

CANCER

Leukemia

We selected leukemia as the primary health endpoint for our health benefits analysis. Significantly increased risks of leukemia have been consistently reported in benzene-exposed workers of various industries, leading EPA to classify inhaled benzene as a “known/likely” human carcinogen under its 2005 cancer guidelines (USEPA, 2005b). In the EPA document *Carcinogenic Effects of Benzene: An Update* (USEPA, 1998), it states “[e]pidemiologic studies and case studies provide clear evidence of a causal association between exposure to benzene and leukemia” (page 4).

Two groups of benzene-exposed workers have been extensively studied and peer-reviewed. The first consists of a group of 1,717 white male workers employed between 1940 and 1972 in Pliofilm manufacturing plants located in Ohio (hereafter, the "Pliofilm Cohort").[21] The second consists of nearly 75,000 workers in a variety of industries in China employed between 1972 and 1987 (hereafter, the "Chinese Worker Cohort"). Results from retrospective analyses of these workers indicate an association between exposure to a range of benzene concentrations and an elevated risk of leukemia (all types). Recent analyses comparing exposed workers to unexposed workers in the

---

[21] Pliofilm is a glossy membrane made from rubber hydrochloride and used chiefly for water-resistant materials and packaging (Crump, 1994).

Chinese Worker Cohort show that exposed workers were roughly two and a half times more likely to develop leukemia than the unexposed workers (Yin et al., 1996, Hayes et al., 1997). Pliofilm Cohort analyses have found similar results comparing the observed cases of leukemia in the cohort to an expected number of cases based on US sex- and age-specific rates (Crump 1994, 1996; Rinsky, 2002). Appendix C provides information on other recently published epidemiologic studies that have found an overall increase in risk of leukemia (all types) with exposure to benzene, or a trend of increasing relative risks (RRs) with increased exposure to benzene (Ireland et al., 1997; Costantini et al., 2003; Adegoke et al., 2003; Sorahan et al., 2005; Guenel et al., 2002; Bloemen et al., 2004; Glass et al., 2003; Collins et al., 2003).

There are four subtypes of leukemia: Acute Myelogenous Leukemia (AML), Acute Lymphocytic Leukemia (ALL), Chronic Myelogenous Leukemia (CML), and Chronic Lymphocytic Leukemia (CLL). The strength of evidence supporting a link between benzene and specific types of leukemia varies. AML has the most evidentiary support for a link with benzene exposures, including associations found in both of the major cohort studies.[22] However, other recent studies identified through the literature search have only found non-significantly elevated risks of AML with benzene exposure or suffer from methodological limitations, such as small numbers of cases or possible exposure misclassification, making the results difficult to interpret.

Based on evidence gathered by EPA in the Integrated Risk Information System (IRIS) support document for benzene carcinogenicity as well as the results of the literature review on the health effects of benzene exposure performed by IEc, we chose to quantify the avoided cases of leukemia due to changes in benzene exposure through a dose-response analysis. We decided to use the outcome of all leukemias for the primary estimate, since this endpoint is the most data rich, compared to the limited evidence for a link with benzene and the specific leukemia types (i.e., AML, ALL, CML and CLL). However, because AML was the subtype with the most evidentiary support, we performed a sensitivity analysis to estimate the number of avoided cases of AML.

Other Cancers

In addition to leukemia, benzene exposure has been associated with other cancerous health endpoints in epidemiologic studies, such as Hodgkin's Lymphoma (HL) and non-Hodgkin's Lymphoma (NHL) (Hayes et al., 1997), multiple myeloma (Rinsky et al., 1987 & 2002; Wong et al., 1995), and myeodysplastic syndrome (MDS) (Hayes et al., 1997). However, data on these endpoints are inconsistent and do not yet support a quantitative evaluation of avoided cases due to benzene exposure.

---

[22] The Chinese Worker Cohort found an elevated RR of acute non-lymphocytic leukemia (ANLL) incidence of 3.0 (95% CI: 1.0, 8.9) and 3.1 (95% CI: 1.2, 10.7) (Hayes et al., 1997; Yin et al., 1996) and the Pliofilm Cohort identified a RR of AML deaths of 5.03 (95% CI: 1.84, 10.97) (Wong, 1995).

NON-CANCER

Benzene exposure at high concentrations has been associated with various hematological abnormalities, including aplastic anemia; however, these adverse non-cancer health effects are unlikely to occur at levels expected to be found in ambient air (less than 10 μg/m^3, based on EPA's NATA study).

EPA developed a chronic reference concentration (RfC) of 0.03 mg/m^3, based on decreases in lymphocytes (a type of white blood cell) reported in a cross-sectional study of a subset of the Chinese Worker Cohort (Rothman et al. 1996a).[23] This study found blood cell effects at exposure concentrations of about 8 parts per million (ppm).[24] The IRIS profile states that decreased lymphocyte count is a biomarker of exposure and is also thought to have a potential role as a "sentinel" effect (i.e., an early sign of toxicity in the bone marrow), but the effect itself is of uncertain clinical significance to the average population (USEPA, 2007b). The significance of the effect depends both on the magnitude of the decrease in lymphocytes and an individual's baseline lymphocyte level.[25]

2.4.2 DOSE-RESPONSE EVALUATION

The following section describes our evaluation of the existing epidemiologic evidence examining the link between benzene and leukemia and how that informed our selection of a leukemia dose-response function for our health benefits model. Specifically, it describes the major cohort studies, the shape of the dose-response relationship, and the cessation lag, which is the estimate of how quickly cancer risks in a population will decline to a new steady-state level following a reduction in exposure.

Choice of Epidemiologic Data

EPA's IRIS identifies the Pliofilm cohort results as the best available data for dose-response evaluation (Rinsky et al., 1981, 1987). IRIS reports a range of inhalation unit risk (IUR) estimates for benzene-induced leukemia (2.2×10^{-6} to 7.8×10^{-6} per μg/m^3 benzene in air; USEPA, 1998) based on a reanalysis of the Pliofilm Cohort data by Crump (1994).[26]

---

[23] An RfC is an estimate (with uncertainty spanning perhaps an order of magnitude) of a continuous inhalation exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime (USEPA, 2007c).

[24] More recent epidemiological and animal studies have found decreased lymphocyte counts at lower exposure levels (Turtletaub and Mani, 2003; Lan et al., 2004; Qu et al., 2002).

[25] For example, the effect of reduced lymphocytes might be more significant for individuals whose immune systems were compromised (e.g., those suffering from HIV/AIDS).

[26] An IUR represents the excess lifetime cancer risk estimated to result from continuous exposure to an agent at a concentration of 1 μg/m^3 in air. While these are typically upper bound estimates, the range of IUR estimates reported for benzene are best statistical estimates derived from observable dose responses using a linear extrapolation model to estimate low environmental exposure risks (USEPA, 2007c).

Strengths of this cohort study include the lack of confounding by other chemicals since workers were exposed primarily to benzene and it is likely that increased risks found in these analyses were due to benzene exclusively; the exposure experienced by this cohort has a wide range, consisting of both high and low exposures; and two sets of exposure estimates were used, Crump and Allen (1984), and Paustenbach et al. (1992), providing a range of estimates. However, the Pliofilm cohort has relatively small number of leukemia cases (14) and some uncertainty in the exposure estimates because there is limited monitoring data in the early years of the study (before 1946).

After reviewing the analytical plan for this case study presented by IEc in 2003, the Science Advisory Board (SAB) Health Effects Subcommittee (HES) recommended that EPA take a closer look at the Chinese Worker Cohort analysis as well as other available data as a possible replacement for the risk estimates of Crump, since this study includes a small number of cases of leukemia in its analysis (USEPA, 2004e).

We found the Chinese Worker Cohort to have a much larger sample size and number of cases (47) than the Pliofilm Cohort. In addition, researchers of this cohort found positive associations between benzene and leukemia at levels closer to ambient (e.g., for workers with <10 ppm average exposure). However, limitations of this study include possible confounding by occupational exposures to chemicals other than benzene and potential problems with exposure assessment, in that only 38 percent of exposures were based on actual measurements (Dosemeci et al., 1994).

In addition to the Pliofilm and Chinese Worker Cohort studies, we examined a number of cohort and case-control studies linking benzene and leukemia, including two studies of petroleum workers known to have low exposures specifically mentioned by the SAB HES in their recommendations (Rushton and Romanieuk (1997) and Schnatter (1996)). We found that these studies suffer from a variety of methodological weaknesses, such as small cohort size, insufficient exposure assessment, and potential confounding of other exposures that limit the usefulness of these studies for our analysis.

For the purposes of our analysis, we ultimately chose to use dose-response slope factors reported by Crump (1994) for our primary estimate of avoided leukemias because the IRIS profile for benzene currently supports the use of data from the Pliofilm cohort for calculating potency estimates. However, despite its limitations, the Chinese Worker Cohort data has certain advantages over the Pliofilm Cohort, such as large sample size and benzene exposure levels that are more consistent with ambient exposures. Therefore, we performed a sensitivity analysis using the results of the Chinese Worker Cohort.

Shape of the Dose-Response Relationship

The shape of the dose-response function for leukemia and benzene is uncertain, with different studies suggesting one or more possible functional forms in the observable range (e.g., linear, supralinear). This makes extrapolating the dose-response function to low levels, such as those found in this study, uncertain as well. Linear models in the observable range were found to be the best fit in the Crump (1994) analysis of the

Pliofilm Cohort.[27,28] The author concluded that "[t]here was no indication of either [cumulative exposure]-dependent or intensity-dependent nonlinearity in the dose responses for any model based on the Crump and Allen exposure matrix" (Crump, 1994, page 234).[29] We also found evidence supporting a supralinear dose-response relationship between observed benzene concentrations and leukemia. For example, an analysis of the Chinese Worker Cohort found that effect estimates tended to plateau at higher levels of benzene (Hayes et al., 1997).[30]

In addition, conflicting information exists regarding the possibility of a threshold in the dose-response function. In our literature search, we found some evidence of a potential threshold in that statistically significant increases in leukemia are not seen at lower exposures levels in the Pliofilm Cohort studies. However, these analyses are uncertain due to minimal statistical power at low benzene levels (see Appendix C for more information).

We chose to use a linear model throughout the range of exposure concentrations in our analysis for several reasons.[31] First, we did not find current evidence on potential thresholds for benzene-induced leukemia to be persuasive. Furthermore, the best fitting models from our chosen epidemiological dataset, the Pliofilm Cohort, were linear in the observable range. Finally, EPA (1998) concludes that "[t]oo many questions remain about the mode of action for benzene-induced leukemia for the shape of the dose-response function to be known with certainty" (page 34). According to EPA's *Guidelines for Carcinogen Risk Assessment*, linear extrapolation to low doses should be used when there is insufficient data to establish a mode of action (MOA) as a default approach because linear extrapolation "generally is considered to be a health-protective approach" (USEPA, 2005b, page 3-21).

---

[27] Specifically, linear multiplicative risk models, where the leukemia mortality rate is proportional to both the change in exposure and the baseline rate of dying from leukemia, were the best fit.

[28] Crump (1994) did not investigate supralinear models; the linear model was the best fit when compared to sublinear models.

[29] We explored the possibility of performing a sensitivity analysis using an intensity-dependent quadratic function reported in the Crump (1994) analysis. In this case, the intensity of the exposure was given greater weight than the duration of exposure. Only borderline significant results were found for this model, using the Paustenbach exposure estimates. Ultimately, we decided not to include this sensitivity analysis because it would have required substantial revisions to the life-table model's exposure processing routing, which we felt were not justified, given the borderline significance of fit of this model.

[30] Additional Chinese Worker Cohort analyses found that benzene metabolite levels plateau at higher benzene exposures, potentially suggesting the existence of an enzyme-mediated process for benzene toxicity that could involve saturation of the enzyme at higher doses (Rothman et al., 1996b & 1997).

[31] We selected a linear dose-response relationship assumption for the observable range as well as for extrapolation to low doses.

Cessation Lag

The term "cessation lag" refers to the estimate of how quickly cancer risks in a population will decline to a new steady-state level following a reduction in exposure.[32] In the original analytical plan, we proposed to use a five-year "cessation lag" for benzene-induced leukemias. The SAB HES, in their review of the report, suggested that we re-examine whether our lag approach was consistent with the epidemiologic literature on this subject.

The ideal data for modeling cessation lag would come from studies that follow the pattern of changes in risk in a study population over time following an exposure reduction. Where such data are limited or unavailable, information on the distribution of latency in a population can be useful for bounding potential cessation lag periods, because it indicates a period of time over which latent cases of disease at the time of the exposure change may continue to be diagnosed, while the population risk moves to a new steady-state level. During our review, we found only one study that explicitly modeled the cessation lag concept, using an analysis stratified on time since last exposure (Silver et al., 2002). This study found that exposures five to ten years prior to the cessation of exposure have the maximum impact on risk, and that exposures between ten and 15 years prior to cutoff may also contribute to a lesser degree. All of the other studies we reviewed included an estimate of latency in their models (i.e., the delay between the critical exposure and manifestation of disease or death). While not the same as the cessation lag, information about latency can also help inform our estimate for a cessation lag. Of the studies examining latency, most found that latency periods of 10 years or less were the best fit for the data. A few found latency periods as long as 15 years.

Rather than incorporating a cessation lag into the benefits as a post-processing step, as EPA has done with other pollutants, such as fine particles ($PM_{2.5}$), we instead chose to select a dose-response slope factor from the Crump analysis that directly incorporates assumptions about the differential impacts of past exposures on current risks. See the section entitled "Incorporating Weighted Exposure" in Section 2.4.3 for more information.

2.4.3 RISK MODEL

Overview of the Model

The purpose of the risk model is to calculate the expected number of fatal and non-fatal cases of benzene-induced leukemia avoided as a result of the implementation of the 1990 CAAA regulations affecting benzene emissions in three counties in the Houston area (Brazoria, Galveston, and Harris). The approach used to estimate these benefits is based on the model used to estimate risks due to radon exposure in the National Research Council's (NRCs) Biological Effects of Ionizing Radiation (BEIR) IV report (1988). The

---

[32] See *Arsenic Rule Benefits Analysis: An SAB Review* Science Advisory Board. EPA-SAB-EC-01-008, August 30, 2001 (USEPA, 2001a) for more information about the concept of cessation lag.

approach consists of a life-table analysis that calculates the probability of contracting (or dying from) leukemia for a given age cohort in a given time period, conditional on the probability of surviving to that period. Figure 2 provides an overview of the life-table model, including the inputs and calculations.

The model first takes the difference between the *With-CAAA* and *Without-CAAA* benzene exposure estimates in each study year to calculate the CAAA-related exposure changes in each year. It calculates these changes for each census tract in the study area and for every five-year age group (e.g., 65 to 69 year olds) residing in each census tract. The model then takes these data to construct a cumulative exposure history for each age group in each census tract, reflecting a total impact of benzene changes, both current and past. Thus, for example, in 2020, the model would calculate an exposure history based on exposure changes experienced from 2020 all the way back to 1990 (or birth, whichever was more recent). Because studies of benzene-exposed workers suggest that exposures may have different effects on leukemia risk depending on when they occur, the model incorporates this information when calculating a cumulative exposure value. For example, in our main model, exposures occurring roughly five years in the past are the most influential for developing leukemia in the current year. Therefore, exposures in that year are given the most weight and exposures occurring before or after that year are given less.

For each census tract, the model uses this cumulative weighted exposure estimate to determine changes in each age group's risk of dying from leukemia in the current period. This size of this effect depends on the size of the cumulative change in benzene exposure; the size of the mortality risks (both all-cause and leukemia-specific) faced by each age group in the baseline; and the assumed relationship (derived from studies of worker exposures) between changes in benzene exposure and changes in leukemia risk. (See equations in Appendix D for further detail). Changes in risk for each age group are then multiplied by the population in that age range in that census tract in that year to calculate the number of avoided leukemia deaths expected to occur among that group in the current year. The model then repeats this process until avoided leukemia deaths are calculated for each five-year age group, census tract, and study year combination. Once completed, the model sums across all age groups and census tracts to produce estimated reductions in avoided leukemia deaths in each county in each study year, and also sums avoided deaths across the whole study area and study period. The final step in the model is to apply an economic value to the avoided leukemia deaths, which is discussed in Section 2.5.

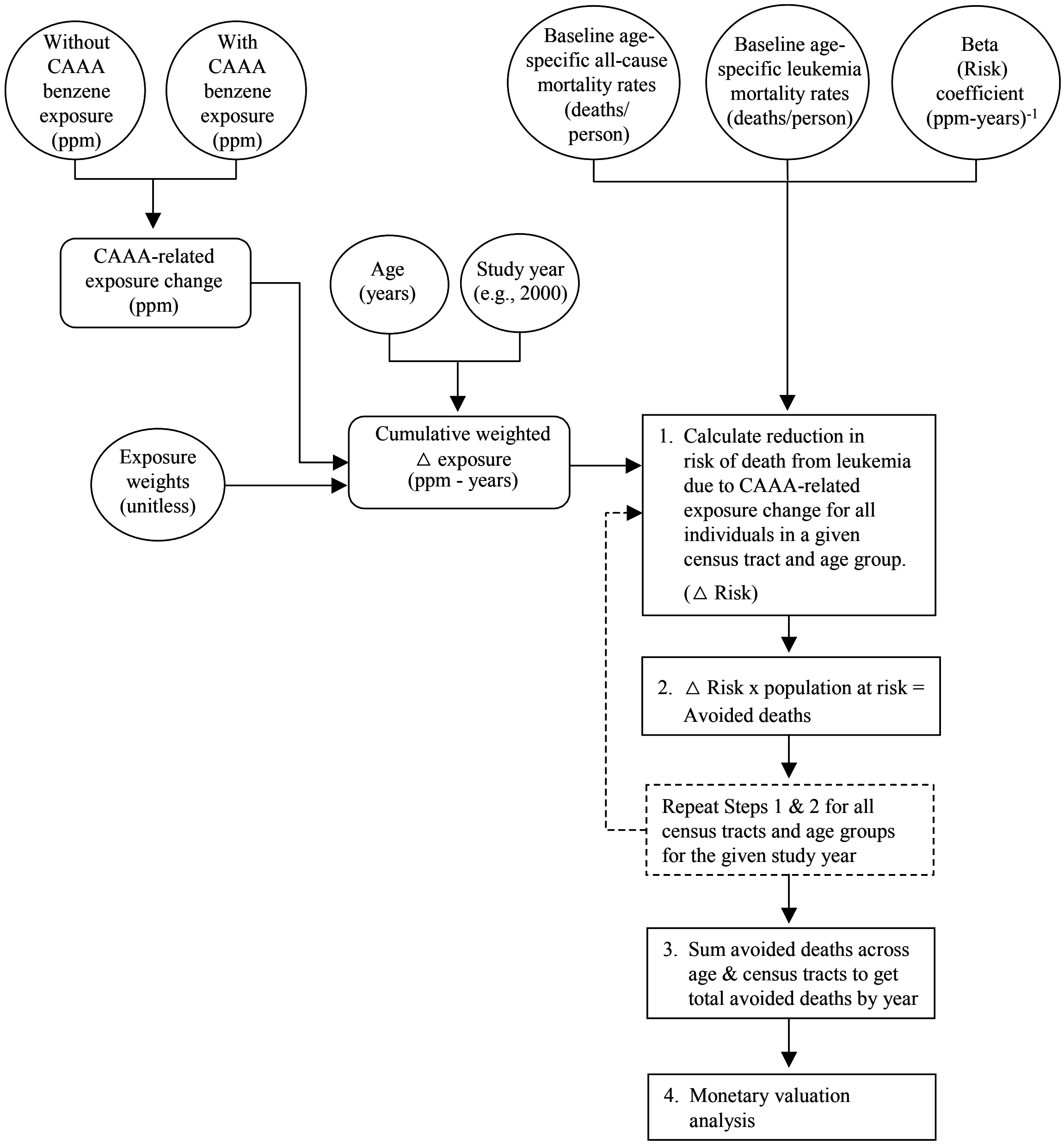
The life-table approach allowed us to estimate benefits to age-specific cohorts, taking into account age-specific mortality rates, both for all-causes and leukemia. This approach also allowed us to explicitly integrate an expected latency period into our model. That is, exposures that were expected to be responsible for initiating the development of leukemia were weighted more heavily and exposures occurring after initiation were weighted less.[33] This approach allows us to estimate a delay in the realization of benefits, but it is

---

[33] This process assumes that once the benzene-induced cancer has been initiated, the time from that occurrence until the resulting mortality is benzene-independent (Crump, 1994).

not necessarily the same as the "cessation lag" effect previously cited by the SAB (USEPA, 2001a). The "cessation lag" refers to the estimate of how fast cancer risks in a population will decline to a new steady-state level following a reduction in exposure. The latency represents the period before any benefits begin to be observed. However, this may provide a reasonable approximation of the cessation lag. See the section entitled "Incorporating Weighted Exposure" below for more information.

FIGURE 2: LIFE-TABLE MODEL OVERVIEW



Note: This flowchart assumes the model is being run with leukemia mortality data. The model can also be run with leukemia incidence data. The difference between the model results for these two runs represents an estimate of avoided non-fatal cases of leukemia.

We calculated a partial lifetime risk of dying from leukemia, focusing on the study period. We estimated the risk change due to the difference in exposure between the *With-* and *Without-CAAA* exposure scenarios for five-year age cohorts at the census tract level.[34] The basic risk equation we used for calculating the partial lifetime probability of dying from leukemia (R) is below. (See Appendix D for a more in depth description of the risk model, including more detailed exposure and risk equations.)

$$R = h/h^* \times S \times (1-q)$$

Where: R = risk of dying from leukemia in the current year, given survival up to that year;

h = leukemia mortality rate;

h* = all-cause mortality rate;

S = probability of surviving up to the current year;

q = probability of surviving through the current year; and

1-q = probability of dying during the current year.

We then multiplied these partial lifetime probabilities of leukemia by the population of the specific age cohort in that census tract in that year to estimate the number of avoided cases.

Survival rates for leukemia have improved since the time of the Pliofilm cohort, suggesting that an increased percentage of leukemia incidence in the study period (1990-2020) will be non-fatal. Non-fatal leukemia cases represent a separate health endpoint in our benefits analysis. Therefore, we ran the risk model using both leukemia mortality and incidence rates with the same dose-response slope factor. The difference between these results represents our estimate of avoided non-fatal cases of leukemia.[35]

Model Inputs

This section describes the various sources of data that were used in the model. Because the model required large amounts of data, we used a Microsoft Access™ database to perform all calculations. Each of these datasets were constructed in Microsoft Excel™ spreadsheets and uploaded in the Access database. In some cases, we manipulated the original data so that it would be consistent with the parameters of our model. For instance, for mortality and incidence rates, if age cohorts reported in the original data differed from those in our model, we calculated weighted average rates for the model age cohorts, using population data from the same year or years as the rate data.

---

[34] The age cohorts started at 0-4 and ended with 95-99.

[35] Note that we compared the resulting split between fatal and non-fatal cases of leukemia against 10-year survival rates for 1988-2004 from the from the Surveillance Epidemiology and End Results (SEER) website to ensure that our methodology was reasonable (http://seer.cancer.gov/). See Section 3.3.1 for further information.

Population Data

For study years between 1990 and 2000, we used population data from the 2000 US Census.[36] For the remaining study years, we used Woods and Poole population projections (2001), consistent with the main 812 analysis. We extracted the relevant population projection data from EPA's BenMAP model at the census tract level by single year of age. We then aggregated the data to match the five-year age intervals in our life-table model.

Health Data

We acquired county level all-cause background mortality rates from the Texas Department of State Health Services, Center for Health Statistics.[37] We used data from the year 1990, which was the base year of the study period. We procured background leukemia mortality and incidence data from the Texas Department of State Health Services, Cancer Epidemiology and Surveillance Branch, Texas Cancer Registry.[38] These were both county-level rates and were only available as an average over several years (1990-2003 for mortality and 1999-2003 for incidence), due to small numbers of cases in each county.[39]

Exposure Data

The HAPEM6 median concentration representing "total" exposure was used for each age group in each census tract as the exposure values for our primary estimate.[40] However, we first adjusted the HAPEM6 output so that it would be consistent with the structure of our life-table model. For instance, our model assessed risk for 20 five-year age cohorts (e.g., 0-4, 5-9, 10-14), whereas the HAPEM6 output contained only six age groups of differing lengths. Therefore, in order to convert the HAPEM6 data to a format consistent with our model, we did one of the following: 1) if a given five-year age cohort was entirely covered by a HAPEM6 age group, we assigned that cohort the exposure concentration for that HAPEM6 age group; or 2) if the five-year age cohort spanned more than one HAPEM6 age group, we calculated a weighted average exposure concentration, based on the number of years spent in each of the HAPEM6 age groups.[41] In addition, HAPEM6 results were only available for the base year (1990) and target years (2000,

---

[36] http://www.census.gov/.

[37] http://www.dshs.state.tx.us/tcr/default.shtm.

[38] Data supplied by Dr. David Risser of the Texas Department of State Health Services.

[39] Since leukemia is a relatively rare disease, in order to get reliable estimates, it is necessary to average over several years of incidence data. It is possible that the later years included migrants, which could introduce uncertainty in the estimates (if these individuals had been exposed to different benzene levels than other residents).

[40] Total exposure consists of a sum of the ambient air concentrations due to the four source sectors (point, non-point, on-road, non-road, and background).

[41] For example, HAPEM6 estimated exposure for 0-1 year olds and 2-4 year olds separately. To calculate exposure for the 0-4 age group in our risk model, we took a weighted average of the two HAPEM6 exposure concentrations, weighting the first concentration with a value of 2 years, the second concentration with a value of 3 years.

2010, 2020). In order to estimate exposure for each year, we linearly interpolated the exposure concentrations between target years. Because our model calculated risk at five-year intervals, we then took an average over the previous five years (e.g., we averaged the concentrations for 1991-1995 for the 1995 concentration).

In order to be consistent with the epidemiological data used in the life-table model, we then converted the output from $\mu g/m^3$ to ppm.[42] In addition, the dose-response slope factor from Crump (1994) is based on occupational exposures. Therefore, we multiplied the output concentration by a conversion factor so that it would be consistent with a typical occupational exposure scenario.[43]

To reduce model computations, we subtracted the exposure concentrations for the *With-CAAA* scenario from the *Without-CAAA* scenario to obtain a "delta exposure" value representing the change in exposure due to the CAAA for each year, age cohort, and census tract. These delta exposure values were then used in the risk calculations, rather than calculating partial lifetime risk of leukemia for each of the two scenarios separately and then subtracting to obtain the difference in risk.

Incorporating Weighted Exposure

EPA's SAB has defined "cessation lag" as the period it takes for risk to decline to a steady state level following a reduction in exposure (USEPA, 2001a). As described in Section 2.4.2, we identified only one epidemiological study specifically estimating the length of the cessation lag. Therefore, in order to develop a temporal stream of benefits, we relied on data that attempts to characterize the latency period (the time between a critical exposure and the development of symptomatic disease or death). We used these data to create a cumulative exposure value in each year for each age group/census tract combination that reflects the differential weighting of past exposures based on their expected importance for developing leukemia, as described below.

Crump (1994) evaluated benzene risk using several models based on data from the Pliofilm cohort. The dose-response models used in the analysis required that a person's prior exposure to benzene be condensed to a single summary measure (Crump, 1994). The author considered two exposure metrics to calculate this single exposure estimate. The first method, "cumulative exposure," employed a "lag," L, and assigned a weight of zero to the last L years of an individual's exposure, assuming that exposures during the most recent L years do not affect mortality rate. The second method, "weighted

---

[42] In order to convert benzene concentrations from $\mu g/m^3$ to ppm, we first converted $\mu g/m^3$ to mg/m^3 by multiplying by 1×10^{-3}. We then multiplied the concentration by 24.45 (a constant) and divided by 78.11 g (the molecular weight of benzene).

[43] The conversion factor consisted of the following: (7 days/week / 5 days/week) × (24 hr/day / 8 hr/day) × ($0.833\ m^3/hr$ / $1.25\ m^3/hr$). The value of $0.833\ m^3/hr$ is an average breathing rate for the general population. Subjects in the occupational studies on which the risk estimates are based are assumed to exhibit a faster breathing rate of $1.25\ m^3/hr$ during an eight-hour work day.

exposure," utilized a weighting function that increases from zero to a maximum for exposures that occur K years in the past, if K represents the best estimate of latency.[44]

A cumulative exposure value can then be calculated for each age cohort in each year in each census tract by weighting previous exposures based on one of the two methods described above and then summing them.

For the purposes of our life-table model, we chose to use the dose-response slope factor from the linear multiplicative model for all leukemia that incorporated the "weighted exposure" method from the Crump analysis. The value of K for this model was 5.3 years. We selected this model because the latency estimate was consistent with the literature, most of which reported latency estimates between 5 and 10 years. In addition, unlike the "cumulative exposure" model, which applies equal weight to all exposures that occur before the latency estimate, the "weighted exposure" model applies lower weights to exposures far in the past, which is more consistent with the literature, where no studies found latency to be greater than 15 years. Because of the uncertainty in the true latency period for benzene-induced leukemia, we performed sensitivity analyses using the "cumulative exposure" model with a five-year lag and a zero-year lag.

Model Output

The model output consisted of the number of deaths from leukemia that were avoided due to the presence of the CAAA for each age cohort in each census tract over a five-year period. We first divided the estimate of avoided deaths by five to obtain an annual avoided deaths value for each year in the study period. We also summed the avoided deaths across all age groups in all census tracts, resulting in an overall cumulative sum of avoided deaths for each county for the entire study period. In addition, we estimated the number of cases expected to occur after the end of the study period that are due to CAAA-related benzene changes within the study period. See the "Expected Total Benefits" subsection of Section 3.3.1 for further information.

2.4.4 ADDRESSING HIGH-EXPOSURE SUBPOPULATIONS

The life-table model used in this case study used median benzene concentrations to estimate avoided cases of leukemia. Therefore, to provide a more complete illustration of the effects of reducing benzene exposures to populations in the Houston area, we performed supplemental calculations of risk reductions to three high-end exposure groups: residents living in census tracts with the highest benzene exposures, residents living near major roadways, and residents with attached garages.

Residents Living In Census Tracts With High Exposure

As part of our assessment of highly exposed subpopulations, we examined CAAA-related reductions in the risk of leukemia for individuals living in census tracts with the highest

---

[44] The weighting function took on the following form: $w(t) = (t/K^2) \exp(-t/K)$. Where: t = the number of years prior to the current year; and K = number of years prior to the current year when the weight reaches its maximum (this also represents the latency estimate).

levels of benzene. We first selected the two tracts in each of the three counties included in this case study with the highest exposure concentrations from HAPEM6 under the *Without-CAAA* scenario in 2020. These tracts also exhibited the highest changes in exposure between the *With-* and *Without-CAAA* scenarios. We then calculated an estimate of individual lifetime risk of leukemia in each of the six tracts under both the *With-* and *Without-CAAA* scenarios, assuming continuous lifetime exposure to median 2020 levels, using the following equation:

$$\text{Individual Lifetime Risk of Leukemia} = \text{EC} \times \text{IUR}$$

Where: EC = median 2020 exposure concentration from HAPEM6 ($\mu g/m^3$); and

IUR = benzene inhalation unit risk estimate from IRIS ($(\mu g/m^3)^{-1}$).[45]

We then subtracted the individual lifetime risks of leukemia under the *Without-CAAA* scenario from the *With-CAAA* scenario to estimate the CAAA-related risk reduction for each of the six tracts.

Note that this assessment of risk is different than the method employed by our main life table model and therefore, the results are not directly comparable. In the calculations above, we are assessing lifetime risk of leukemia, assuming constant lifetime exposure of an individual born today to median 2020 levels of benzene exposure under the *With-CAAA* and *Without-CAAA* scenarios. The life table model calculates partial lifetime risks of leukemia retrospectively over the study period from estimates of cumulative weighted previous exposures.

Residents Living Near Roadways

Another highly exposed subpopulation in the study area includes individuals living in close proximity to major roadways, such as the interstate loops in Harris County that surround downtown Houston. A substantial number of studies have demonstrated increased concentrations of benzene and other mobile source related HAPs near roadways. For example, Kwon (2005) analyzed ambient VOC measurements in Elizabeth, New Jersey from the Relationship among Indoor, Outdoor, and Personal Air (RIOPA) study and found that ambient concentrations of benzene, toluene, ethylbenzene, and xylenes measured near homes within 200 m of roadways are 1.5 to 4 times higher than urban background levels. Several other studies have found that concentrations of benzene and other mobile source air toxics are significantly elevated near busy roads compared to "urban background" concentrations measured at a fixed site (e.g., Skov et al., 2001; Jo et al., 2003; Fischer et al., 2000; Ilgen et al., 2001; and Sapkota et al., 2003).

Version 6 of the HAPEM model, which was applied in this study, includes algorithms that account for the gradient in concentrations of primary (directly emitted) mobile source air toxics within 200 meters of major roadways (ICF, 2007). HAPEM6 adjusts ambient concentrations generated by AERMOD for each census tract using concentration

---

[45] Note that the IRIS profile for benzene presents a range of values for the IUR (2.2×10^{-6} – 7.8×10^{-6}). We calculated values using both ends of the range.

gradients developed with the CALPUFF dispersion model (Cohen et al., 2005). For locations within 75 meters and from 75 to 200 meters from major roads, HAPEM6 adjusts ambient concentrations upward, while concentrations at locations further from major roadways are adjusted downward. These adjustments are consistent with results from prior modeling studies that explicitly accounted for concentration gradients around major roads within census tracts (Cohen et al., 2005; Stein et al., 2007). HAPEM6 then applies the adjusted concentrations in its microenvironmental concentration calculations.

To assess the impact of HAPEM's near-roadway algorithm on our primary results and on exposures to highly exposed individuals, we conducted an additional HAPEM run for 2020, turning off the near-roadway feature. We then compared the difference in the annual average benzene concentration in 2020 between the *With*- and *Without-CAAA* scenarios for these two HAPEM runs.

Residents With Attached Garages

We also performed supplemental calculations of risk reductions to residents living in homes with attached garages. Studies of benzene levels in homes with attached garages suggest that residents in these homes may be exposed to higher indoor benzene concentrations than residents in other types of homes (Gordon et al., 1999; Schlapia and Morris, 1998). While the population living in homes with attached garages may benefit from reductions in benzene emissions that occur in-garage, we were unable to identify data on local benzene concentrations in attached garages in the Houston area with which to estimate those benefits. Therefore, we performed an illustrative, back-of-the-envelope calculation to assess the rough magnitude of additional potential benefits that may result from CAAA-related reductions of in-garage benzene emissions in 2020. Appendix E contains a detailed description of the calculations we performed, including the equations used. We provide a brief overview of the process below.

Our approach involved the following three steps:

1) We assessed the CAAA-related percent reduction in total emissions in the non-road and on-road categories that are expected to occur within attached garages in 2020. The percent reduction was based on the difference in the in-garage emissions between the *With*- and *Without-CAAA* scenarios in 2020. We used slightly different approaches for determining the non-road and on-road portions of the total emissions under the *With*- and *Without-CAAA* scenarios, due to the available data for each of these categories (See Appendix E for more information);
2) We applied the percent reduction in emissions to an estimate of average benzene exposure attributable to attached garages based on previous U.S. studies;[46] and

---

[46] We used an estimate of average indoor benzene exposure attributable to attached garages from Appendix 3A of the MSAT Regulatory Impact Analysis (RIA) (USEPA, 2007d). We selected a value that excluded studies conducted in Alaska, due to a number of differences expected in attached garage-related exposures between Alaska and Houston (see Appendix E for additional information).

3) We calculated the annual number of avoided cases of leukemia in the Houston area in 2020 that would be expected based on the CAAA-related reduction in attached garage-related exposures using the value calculated in step 2, the benzene IUR from IRIS, and an estimate of the size of the affected population.

2.4.5 ESTIMATING NON-CANCER HEALTH EFFECTS

We considered extrapolating the dose-response function based on the data supporting the RfC in order to estimate "cases" of reduced lymphocyte counts expected at environmental exposure levels. However, the data set supporting the RfC is limited (2 data points) and does not support an extrapolation beyond the benchmark concentration down to ambient levels. We identified in our 2005 literature search two additional studies linking reduced lymphocyte count to occupational benzene exposure, both of which had lower exposure concentrations (below 1 ppm) and larger number of data points (3 and 4). Ultimately, extrapolating these studies to low doses proved to be too time and resource intensive for this case study. Therefore, we assessed this endpoint using the approach outlined in the original analytic plan (IEc, 2003), reporting the difference in the number of individuals experiencing benzene concentrations above the RfC under the *With-CAAA* and *Without-CAAA* scenarios.

2.5 BENEFIT VALUATION

This section describes our approach to assigning economic value to the estimated benefits of reductions in ambient benzene concentrations. The scope of the valuation methodology is determined by the prior steps in the case study, which necessarily limits monetization to those health effects for which dose-response functions are available. Therefore, other benefits of reductions in benzene likely exist and have value (e.g., non-cancer health effects, cancers other than leukemia), but we were unable to quantify them in the framework of this case study.

2.5.1 OVERVIEW OF APPROACH

We applied valuation methods that are largely consistent with those employed to value the benefits of the Second 812 Prospective analysis of criteria pollutants (see Chapter 8). That analysis employed a Value of Statistical Life (VSL) estimate to assign economic value to avoided deaths from air pollutants. In the benzene exposure case, however, there is the additional consideration of medical costs associated with the period of cancer illness (the morbidity increment) leading up to death (hereafter, "pre-mortality morbidity"). In addition, we have also valued non-fatal cancer cases, which are not reflected in the criteria pollutant analysis. In order to value these non-fatal cancer cases, we followed recent SAB advice on this topic given during a consultation in 2001 regarding the arsenic in drinking water rule-making by EPA's Office of Water, discussed in more detail below.

2.5.2 VALUATION OF CANCER ENDPOINTS

Fatal Cancers

Value of Statistical Life

Fatal cancers were valued on a per-case basis using a VSL estimate presented in a meta-analysis of several U.S. wage-risk studies by Viscusi and Aldy (2003). We used a mean value of $7.4 million at 1990 income levels (in 2006 dollars).[47] We then applied income elasticity values for premature mortality to account for the projected growth in willingness-to-pay (WTP)-based VSL estimates that is associated with real income growth.[48] This results in an adjusted VSL value for each year subsequent to 1990. The resulting VSL estimate for 2020, for example, was $8.9 million in 2006 dollars.

Pre-Mortality Morbidity

For this analysis, in addition to using VSL to estimate the benefits of avoided cancer deaths, we also provide an estimate of the value of avoided morbidity associated with deaths from cancer. The procedure we apply is consistent with EPA SAB advice delivered as part of prior reviews of both a cancer valuation white paper in 2000 and an economic analysis of more stringent standards for arsenic in drinking water.[49]

To summarize the SAB advice, a special panel of the SAB Environmental Economics Advisory Committee (EEAC), in its review of the EPA Office of Ground Water and Drinking Water's (OGWDW) Arsenic in Drinking Water Rule, endorsed adding estimates of the medical costs of treatment and amelioration for fatal cancers to the VSL as a lower bound on the true (total) value of avoiding fatal cancers (USEPA, 2001a).[50] As a preface to this endorsement, the panel had acknowledged that, as a general recommendation, there was insufficient evidence to support a broad "cancer premium" for the avoidance of fatal cancer risk relative to other types of fatal risk reflected in the VSL typically applied

---

[47] This value is from Table 8 of Viscusi and Aldy (2003) and represents the mean predicted VSL for the U.S. sample using Huber Weights (model 5). This estimate was selected because it was the best model for the data, had relatively tight confidence bounds and reduced non-normality in the error term by using Huber weighting. We adjusted the reported value ($6.3 million in 2000 dollars) for inflation to 2006 dollars using the standard inflators reported in BenMAP (USEPA, 2008).

[48] The specific income elasticity values and per-capita income growth estimates combine to yield annual adjustment factors for the growth in WTP over time. The annual adjustment factors were taken from BenMAP (USEPA, 2008a) for all years up to 2024. For years after 2024, we estimated an approximate income adjustment factor growth rate and applied that rate to generate annual adjustment factors through 2030.

[49] See USEPA (2001a). *Arsenic Rule Benefits Analysis: An SAB Review*. Science Advisory Board. EPA-SAB-EC-01-008, August 30, 2001; and USEPA (2000). *An SAB Report on EPA's White Paper Valuing the Benefits of Fatal Cancer Risk Reductions*. Science Advisory Board. EPA-SAB-EEAC-00-013, July 27, 2000.

[50] Note that this specific adjustment has been subsequently applied in several economic analyses supporting final OGWDW rules. It was most recently applied in the *Economic Analysis for the Final Stage 2 Disinfectants and Disinfection Byproducts Rule*, USEPA Office of Water (4606-M), EPA 815-R-05-010, December 2005. See page 6-83 for a brief description of the procedure applied in that RIA, which closely follows the procedure we have used here.

by EPA.[51] Just prior to the issuance of this report, the larger EEAC had also concluded that, while a cancer premium for morbidity, dread, and fear was valid in principle, there was insufficient evidence to apply any specific WTP adjustment to the standard VSL to reflect a cancer premium. This finding was reflected in subsequent SAB review of the benefits of the arsenic rule; that panel, however, did not consider the option of adjusting the VSL to reflect the medical cost of illness for cancers (USEPA, 2000).[52]

A few additional studies since this time have further tested the idea of a cancer premium for VSL. Most recently, Van Houtven, Sullivan, and Dockins (2008), found that WTP to reduce cancer risk with a five year latency period is three times larger than WTP to reduce current automobile-accident risks, although the cancer premium declined as respondents considered cancers with longer latency periods. In addition, Hammitt and Liu (2004) found respondents in a Taiwanese stated preference survey were willing to pay about 30 percent more to reduce their risk of contracting a fatal cancer versus a similar non-cancer illness, though the results were only weakly significant. Finally, Tsuge, Kishimoto, and Takeuchi (2005) identified a small but significant preference for avoiding cancer risks in Japan. While this literature is growing, we believe it is premature at this time to develop risk-attribute-based adjustment factors for VSLs that specifically address WTP to reduce cancer mortality risks, and instead apply only the pre-

---

[51] The full quote from EPA, 2001a reads as follows, "We believe that the central estimate of $6.1 million for the value of a statistical life (VSL) is appropriate. On the question of whether to add a value for cancer morbidity before death, we do not believe that there is an adequate basis in the literature for doing this. But we can endorse adding estimates of the medical costs of treatment and amelioration for fatal cancers to the VSL as a lower bound on the true value of avoiding fatal cancers" (from page 5-6 in the referenced SAB report).

[52] The full quote from USEPA (2000) reads as follows, "The Committee supports the principle that the morbidity, fear, or dread associated with cancer is a valid component of the cost that individuals attribute to the incidence of cancer. Thus, in principle, the value of reductions in cancer risks should include both the value of the reduced risk of death and the value of reduced risk of the morbidity, fear, and dread that precedes the death incident. To the extent that cancer victims typically suffer greater morbidity, fear, or dread than the victims of the causes of death involved in VSL studies, it would be appropriate to attach a “cancer premium” to the value of an avoided death from cancer. The Committee finds, however, that existing studies provide little reliable information as to the magnitude of this premium, and concludes that until better information becomes available, it is best not to assign such a premium.

The white paper cites studies by Savage (1993) and by Jones-Lee, Hammerton, and Philips (1985) as evidence that people are willing to pay a “cancer premium” to avoid fatal cancers relative to other fatal risks. The paper cites a suggestion from Revesz (1999) that the VSL for an immediate fatality be adjusted by “at least a factor of two” to capture the morbidity, fear, and dread associated with cancer.

The Committee disagrees with this suggestion for two reasons. First, the articles by Savage and Jones-Lee et al. do not measure an individual’s willingness-to-pay (WTP) to avoid fatal cancer; hence they cannot be used to justify the proposed adjustment. Jones-Lee et al. ask respondents if they could reduce deaths from one of three causes — motor accidents, heart disease and cancer — by 100 persons annually, which cause would they select. The respondent is then asked how much he or she would pay for this reduction. This question measures WTP to reduce risks to others as well as to oneself, whereas the VSL values private risk reductions. Similarly, the Savage article does not elicit private WTP but asks the respondent to allocate $100 among “commercial airplane accident research,” “household fires research,” “automobile accident research,” and “stomach cancer research.” Second, the appropriate way to determine whether a “cancer premium” is required is to value reductions in the risk of a fatal cancer directly. There is only one study (Magat et al. 1996) that has attempted to value reductions in fatal cancer risk directly. For the case of fatal lymphomas it suggests that no cancer premium is warranted. Clearly, further research is called for in this area. The Committee believes that until empirical work clearly establishes the value of this premium, it is best not to attempt to apply one " (from page 5-6 of the referenced SAB report).

mortality morbidity adjustment described above. This approach is consistent with recent advice from the SAB Council provided as part of its review of this case study (USEPA, 2008b).

Based on the available literature and SAB advice, we conclude that the VSL applied to value avoided fatal leukemia risks represents the value of avoiding a premature death, but it does not explicitly take into account the medical costs associated with the period of illness (the morbidity increment) leading up to death. Based on estimates presented in EPA's Cost of Illness Handbook (USEPA 1999b) for a "typical" cancer case, we estimate the medical costs for a fatal leukemia case to be $98,971 at 1996 price levels.[53] This cost can be updated to 2006 price levels using the Consumer Price Index (CPI) for Medical Care (see USGPO 2009); the result is $145,810, which for our purposes we round to $150,000 and apply as a point estimate to each fatal case of leukemia in the benefits model.[54]

Non-Fatal Cancers

To our knowledge, EPA's Office of Air and Radiation (OAR) has not previously developed or published an estimate to value non-fatal cancers. In addition, the overall EPA Guidelines for Economic Analysis provides only general guidance on valuation of non-fatal morbidity; in summary, WTP values are preferable, but cost-of-illness values are also acceptable.

EPA's OGWDW, however, has applied existing valuation estimates to non-fatal cancers. Prior to 2001, valuation of non-fatal cancer in OGWDW economic analyses was based on application of a WTP value for chronic bronchitis, based on the assumption that the severity of a chronic but non-fatal cancer case and a case of chronic bronchitis are roughly similar. That approach was reviewed by the SAB EEAC in 2001. At that time, the SAB recommenced that the chronic bronchitis value be supplemented by a value from

---

[53] Estimate derived from EPA's Cost of Illness Handbook, Chapter II.1, entitled, "Introduction to the Costs of Cancer." See Table II.1-4 - Incremental Undiscounted Direct Medical Costs for a Typical Cancer, on page II.1-26. The estimates presented in that table were for a typical case with a 50 percent mortality rate. We adjusted the reported value for the component attributed to terminal care to reflect a certain fatal case. The result is an estimate of $98,970.84 in 1996 dollars. Cost of Illness Handbook available on EPA's website at: http://www.epa.gov/oppt/coi/ (downloaded July 2005).

[54] CPI-Medical Care series taken from Table B-60 in USGPO 2009, *Economic Report of the President*. (http://www.gpoaccess.gov/eop/tables09.html).

the one study that values a non-fatal cancer, Magat et al. (1996).[55] The Magat et al. study is a stated preference, health risk tradeoff study that evaluated the marginal rate of substitution for risks of non-fatal lymphoma and risk of accidental death from a car accident. The resulting risk-risk tradeoff value can be applied to an estimate of the VSL to generate a value for avoiding a statistical case of non-fatal lymphoma.

In this study, we have followed the previous SAB EEAC advice to estimate the value of a non-fatal cancer case using the chronic bronchitis value and a value from the Magat et al. work to bracket a range of possible values. To generate the endpoints of this range, we derived a WTP value for chronic bronchitis from EPA's September 2006 PM NAAQS RIA ($410,000, 2006$, 1990 income levels)[56], and used the VSL for our primary estimate from Viscusi and Aldy (2003) along with the risk-risk ratio estimated by Magat et al. (0.583) to calculate a non-fatal lymphoma value ($4.3 million, 2006$, 1990 income levels).

Based on SAB advice provided during a review of this case study (USEPA, 2008b), we further examined the appropriateness of using estimates of WTP from chronic bronchitis and non-fatal lymphoma to value cases of non-fatal leukemia by comparing the symptoms, severity, duration, and treatment of these illnesses.

Symptom data for these conditions show significant overlap between leukemia and lymphoma (e.g., fever, weight loss, night sweats, fatigue, enlargement of spleen, loss of appetite, and swollen lymph nodes) and to a lesser extent between leukemia and chronic bronchitis (e.g., shortness-of-breath, fatigue, and headaches) (MedLine, 2008).[57]

The duration and treatment of leukemia varies by subtype. For instance, acute leukemia is usually treated immediately with options such as chemotherapy, targeted drug therapy, biological therapy, radiation therapy or stem cell transplant (NCI, 2008). It is possible for acute leukemia to go into remission or even be cured. Chronic leukemia, however,

---

[55] The full text of the SAB advisory states:

> "To value non-fatal bladder cancers, the Agency used a value for avoiding a statistical case of chronic bronchitis obtained by Viscusi, Magat, and Huber (1991). We have two reservations about this. First, this study used a small sample obtained in a shopping mall in North Carolina and thus may not be representative of either the U.S. population as a whole or the population of individuals at risk of bladder cancer. Second, we have no basis for determining that avoiding a case of chronic bronchitis has the same value as avoiding a non-fatal case of bladder cancer.
>
> On this second point, there is one study of willingness to pay to avoid a non-fatal case on one type of cancer. Magat, Viscusi, and Huber estimated the willingness to pay to avoid a case of non-fatal lymphoma to be $3.6 million (Magat, et al. 1996). This value was obtained from a similar shopping mall intercept survey with a substantially larger sample size. So, although the endpoint being valued more nearly corresponds to non-fatal bladder cancer, there is still the question of the representativeness of the sample. We also note that the value obtained is at least 20 times larger than the cost of illness for non-fatal bladder cancer cited in Exhibit 5-10. Thus we do not have a lot of confidence in this number. Therefore, we recommend that the value used in the report and the alternative discussed here be used as bounds in an uncertainty analysis. However, this range should be clearly identified as displaying the two extreme estimates available in the literature so it is not misconstrued as a confidence interval" (from EPA, 2001a, page 5 and subsequent text).

[56] See U.S. EPA 2006, *Regulatory Impact Analysis for the 2006 National Ambient Air Quality Standards for Particle Pollution*, available for download at: http://www.epa.gov/ttn/ecas/ria.html.

[57] We also attempted to locate health-related quality of life (HRQL) scores (such as the health utilities index (HUI) or the EuroQoL EQ-5D) for these specific conditions, but were unsuccessful.

can be asymptomatic and treatment is sometimes delayed. Chronic leukemia can often be controlled but it is difficult to cure (MedLine, 2008).

The duration and treatment of lymphoma also varies by type. Hodgkin's lymphoma is one of the most curable forms for cancer. Treatment generally includes chemotherapy and/or radiation therapy. Non-Hodgkin's lymphoma is more likely to be cured if it is fast growing. Treatments vary depending on the type of non-Hodgkin's lymphoma and include chemotherapy, radiation therapy, and stem cell transplant (Leukemia and Lymphoma Society, 2008). Chronic bronchitis symptoms can be treated with medication and oxygen therapy. However, this disease is long-term and cannot be cured completely (MedLine, 2008).

In conclusion, based on the data we have gathered on symptoms, duration, and treatment of leukemia, lymphoma, and chronic bronchitis, it appears as though employing a range of WTP values is appropriate for non-fatal leukemia. Leukemia includes several subtypes, each of which varies significantly. Acute non-fatal leukemia likely is a better match with non-fatal lymphoma in terms of symptoms, duration, and treatment. However, chronic non-fatal leukemia is unlikely to be cured and could be seen as similar in duration to chronic bronchitis. Therefore, bounding the WTP estimates for non-fatal leukemia with estimates for these two illnesses appears to be reasonable.

As a crosscheck on these values, we estimated cost-of-illness (COI) values for a non-specific case of non-fatal cancer. Our estimates rely on estimates of the direct medical cost of illness from EPA's *Cost-of-Illness Handbook*. Assuming three months of initial treatment followed by 16 years of follow-up treatment, consistent with the median age of diagnosis for all leukemias of 67 and an approximate average life expectancy at 67 of 16 years, we estimated the net present value in 2006$ of the direct medical cost of illness as $116,000 (using a 7 percent discount rate) to $200,000 (using a 3 percent discount rate). As expected, the direct medical costs are significantly less than a comparable WTP estimate - the direct medical costs exclude such factors as lost earnings, implicit value of lost caregiver time, and pain and suffering of the patient over the period of illness.

We also identified a value for a case of non-fatal cancer used by the European Commission countries in their "ExternE" study of the external costs of energy generation. The value of 450,000 (1995 European Currency Units or ECU)[58] was based on a U.S. COI study that included indirect costs of illness in the form of lost wages.[59] Converting to 2006$, we obtain a value of $700,000. The ratio of the WTP used in this case study to this estimate of COI ranges from 0.5 at the low end of the WTP range for non-fatal leukemia to 5.4 at the upper end of the range. A review of studies estimating both WTP and COI for various illnesses in the *Handbook for Non-Cancer Health Effects Valuation* (SPC, 2000) found that WTP/COI ratios ranged from 2 – 31.5 with a median of 3.9 and a

---

[58] The ECU was a currency used by the member states of the European Union (EU) prior to introduction of the euro on January 1, 1999.

[59] See Table 5.2, page 35 in, Common Annexes of the ExternE National Implementation Reports (1998), downloaded 6/1/07 from: http://externe.jrc.es/reports.html.

mean of 8.1 (see Appendix B, Table B-1). Therefore, the estimates used in this case study fall outside this range on the low end for the WTP estimate derived from the chronic bronchitis study but are between the median and mean of values for the non-fatal leukemia WTP.

Cessation Lag

As discussed previously, reduction in exposure to benzene leads to reduction in cancer cases after a period of cessation lag. In economic terms, it is plausible to assume that individuals would prefer avoidance of immediate health effects relative to avoidance of health effects with a delay, suggesting that their WTP to avoid delayed heath effects is affected. Because the underlying VSL estimates are largely for immediately manifest risks of death, the VSL estimate needs to be adjusted to account for the effect of the cessation lag on WTP.

We made this adjustment by discounting the VSL estimate by the period of cessation lag using four alternative discount rates. We used a discount rate of 5 percent for our primary estimate and used discount rates of 0, 3 and 7 percent as sensitivity analyses.

CHAPTER 3 | RESULTS

This chapter presents the results of the emissions, air quality/exposure, and health effects modeling steps in the analytical chain. We present the health benefit results both as avoided cases of leukemia, and as monetized benefits valued as described in the Valuation step in Chapter 2.

3.1 EMISSIONS

Figure 3 illustrates the difference in emissions of benzene in the Houston-Galveston study area in 2000, 2010, and 2020 under the *With-CAAA* and *Without-CAAA* scenarios. Table 2 provides the specific modeled emission estimates by sector. Both exhibits show that the CAAA have resulted in significant benzene emission reductions in the Houston-Galveston study area since 1990. We first discuss the emissions trends under the *With-CAAA* scenario, and then compare the results for the *With-* and *Without-CAAA* scenarios. For additional details concerning these results, please consult Appendix A.

3.1.1 EMISSIONS UNDER THE *WITH-CAAA* SCENARIO

A significant fraction of the reductions in benzene from the CAAA occurred within the first decade following passage of the amendments.[60] Under the CAAA in 2000, total emissions decreased 70 percent from 1990 levels, with the bulk of this reduction occurring in the combined point and non-point sector.[61] For these sources, the benzene emission reductions during this 1990 to 2000 period are largely attributable to Federal maximum achievable control technology (MACT) emission standards, and local VOC measures in the 1-hour ozone attainment plan that required the petrochemical facilities in the area to reduce hazardous air pollutant (HAP) and/or VOC emissions. The chemical manufacturing and petroleum refining industries achieved the most significant benzene emission reductions in these sectors in this period. Mobile sources also exhibit significant reductions in this period, due in part to existing pre-1990 Tier 1 regulations reducing exhaust and evaporative VOC emissions and in part to CAAA-related

---

[60] These results do not include the impact of the 2007 MSAT rule, which was promulgated too late to be included in the *with-CAAA* scenario.

[61] We have chosen to combine point and non-point emissions into a single category because of a discrepancy in the way that the 1990 and 2000 NEIs treat fugitive emissions from the synthetic organic chemical manufacturing industry ("SOCMI fugitives"). The 1990 NEI includes these emissions in the non-point source category while the 2000 NEI reports them as point source emissions. Because we project *Without-CAAA* emissions for point and non-point sources from 1990 NEI data and project *With-CAAA* emissions for these sources from 2002 NEI data, SOCMI fugitives end up categorized differently under the two scenarios. SOCMI fugitive emissions are a significant source of emissions, contributing nearly 2,400 tpy in 1990; therefore, we have combined the two categories to accurately reflect the combined impact of CAAA measures on point and non-point emissions sources in Houston.

reformulated gasoline requirements and inspection and maintenance (I/M) programs in each county.

FIGURE 3: MAJOR, AREA & OTHER, ON-ROAD, AND NON-ROAD EMISSIONS (TONS) FOR EACH YEAR AND SOURCE TYPE

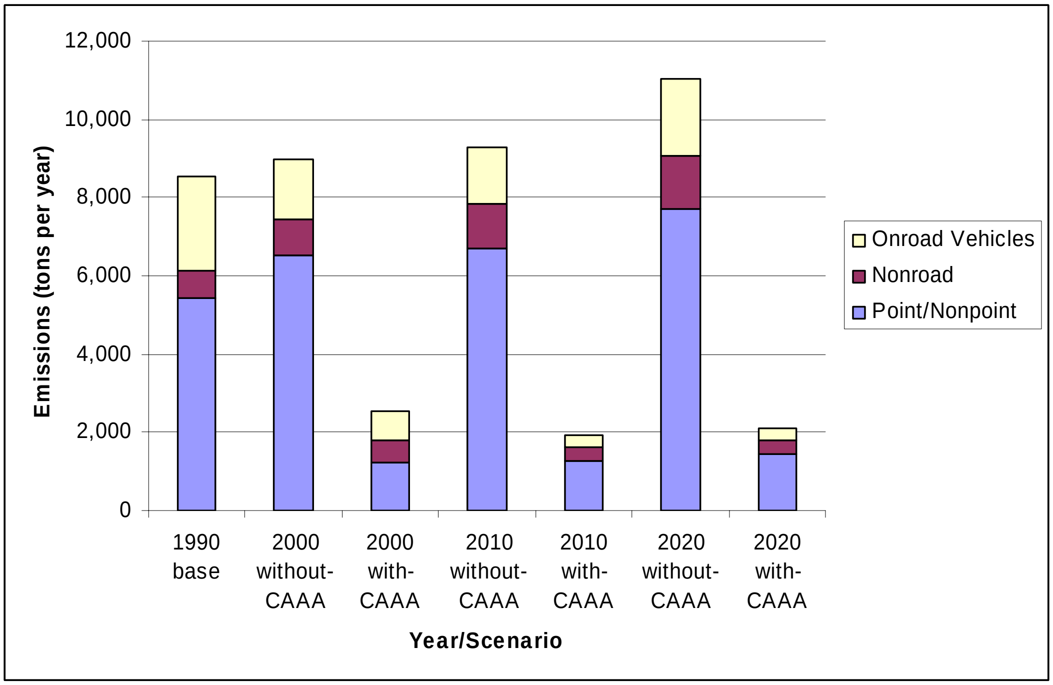


TABLE 2: HOUSTON-GALVESTON BENZENE EMISSIONS SUMMARY (TONS PER YEAR [TPY])

YEAR	1990	2000		2010		2020	
SCENARIO		WITHOUT-CAAA	WITH-CAAA	WITHOUT-CAAA	WITH-CAAA	WITHOUT-CAAA	WITH-CAAA
SECTOR							
Point/Non-point	5,409	6,532	1,230	6,699	1,258	7,702	1,440
Non-road	740	900	567	1,127	354	1,351	360
On-road Vehicles	2,375	1,541	762	1,449	328	1,988	282
Total	8,524	8,973	2,559	9,275	1,940	11,041	2,082

Total emissions continue to decrease in 2010 under the CAAA (77 percent below 1990 levels) and increase slightly between 2010 and 2020. Point and non-point source emissions are essentially stable from 2000 to 2010 and slightly increase from 2000 through 2020. The addition of 7- and 10-year MACT rules in the second decade mitigate emissions growth in this category the first period, but we see a slight increase from 2010 to 2020 in part because the analysis is not applying any new point or non-point source

VOC or benzene-related control programs post-2010.[62] We do observe additional reductions in the mobile source category in this period due to Tier 2 emission standards and associated requirements that lower the sulfur content of gasoline. Reductions in non-road emissions are due largely to the implementation of spark-ignition engine standards.

3.1.2 DIFFERENCE IN EMISSIONS BETWEEN THE *WITH-* AND *WITHOUT-CAAA* SCENARIOS

When we compare the *With-CAAA* scenario to the counterfactual *Without-CAAA* scenario, we observe substantial and increasing differences in each of the three target years – approximately 6,500 fewer tons of benzene in 2000, 7,300 fewer tons in 2010, and nearly 9,000 fewer in 2020. These changes represent reductions in benzene emissions of 71, 79, and 82 percent, respectively, over the *Without-CAAA* scenario. Most of this difference is due to emission controls on point and non-point sources, which emit thousands fewer tons per year under the CAAA; however, reduced emissions from motor vehicles also contribute significantly, particularly in the later years, as the Tier II emissions standards begin to have an impact.[63] Emissions reductions from the non-road sector are a relatively small contributor, because the base year emissions are relatively low; its contribution to overall reductions is greatest in 2010 and 2020.

3.2 AIR QUALITY/EXPOSURE MODELING

The air quality modeling step produced both estimated ambient concentrations of benzene in the study area, using AERMOD, and estimates of age-specific exposure concentrations using EPA's HAPEM that reflect the influence of individuals' activity patterns on the benzene exposure they are likely to experience during their daily activities. Detailed results for both study elements may be found in Appendix B; we provide an overview and comparison of the results from both models below.

Figure 4 summarizes the distribution of benzene concentrations predicted in the study area in the base year and each target year under the *With-* and *Without-CAAA* scenarios. The distributions reflect the variation in concentrations across census block groups in the three counties studied. The yellow *With-CAAA* distributions show both lower median (center line) concentrations under the *Without-CAAA* scenario and tighter distributions with less variation than the green *Without-CAAA* distributions. The difference in medians widens with time, both due to additional CAAA-related benzene decreases (particularly

---

[62] While there may be regulations added in this area in the next few years to meet new nonattainment obligations, based on the current set of Federal and State regulations affecting this area, benzene emission rates for this category have no expected declines in the 2010 to 2020 period other than for woodstoves.

[63] Our model indicates that some benzene emissions reductions from mobile sources occur between 1990 and 2000 even in the absence of the 1990 CAAA, due to fleet turnover enhancing the effects of pre-1990 CAA emissions regulations. As a result, growth in emissions in the first decade of the *Without-CAAA* scenario is less than might be expected, and the percentage reduction in total emissions between the with- and *Without-CAAA* scenarios is not much larger than the percent difference between the 2000 *With-CAAA* scenario and 1990. This effect lessens in 2010 as fewer older cars remain on the road and vehicle miles traveled (VMT) increase. By 2020, the VMT effect dominates and emissions increase in the *Without-CAAA* scenario. Meanwhile, CAAA mobile source provisions such as Tier II emission regulations have an increasing impact from 2000 to 2020, widening the difference between the two scenarios for mobile sources during that time.

between 2000 and 2010) and due to projected emissions growth without the CAAA (particularly between 2010 and 2020).

FIGURE 4: BLOCK GROUP LEVEL TOTAL CONCENTRATION (μg/m^3) DISTRIBUTIONS FOR 1990, 2000, 2010, AND 2020 FOR *WITH-CAAA* (YELLOW) AND *WITHOUT-CAAA* (GREEN) SCENARIOS

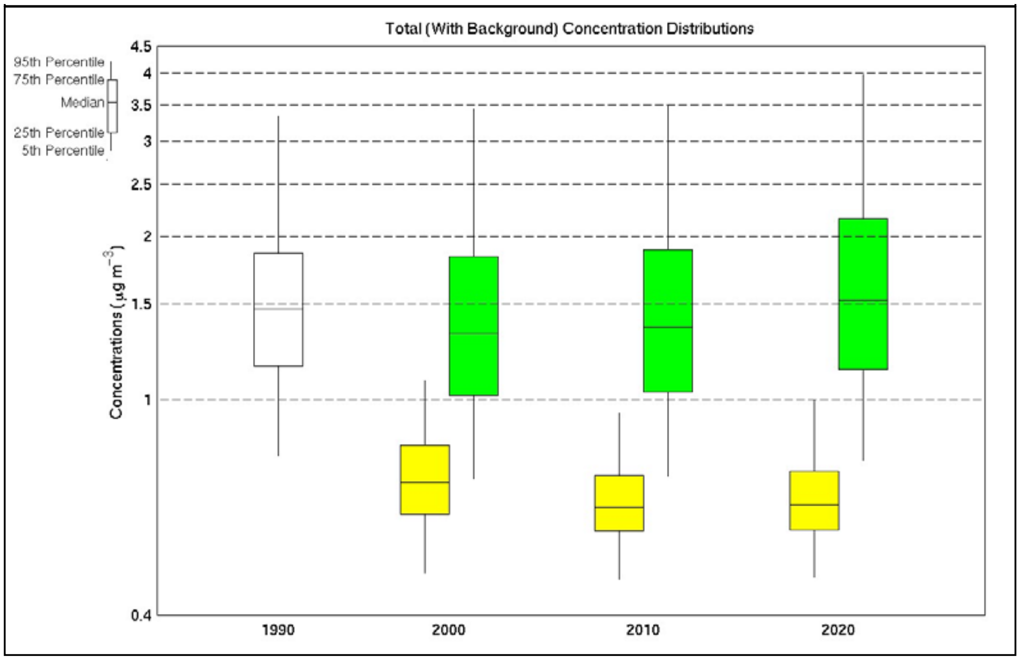
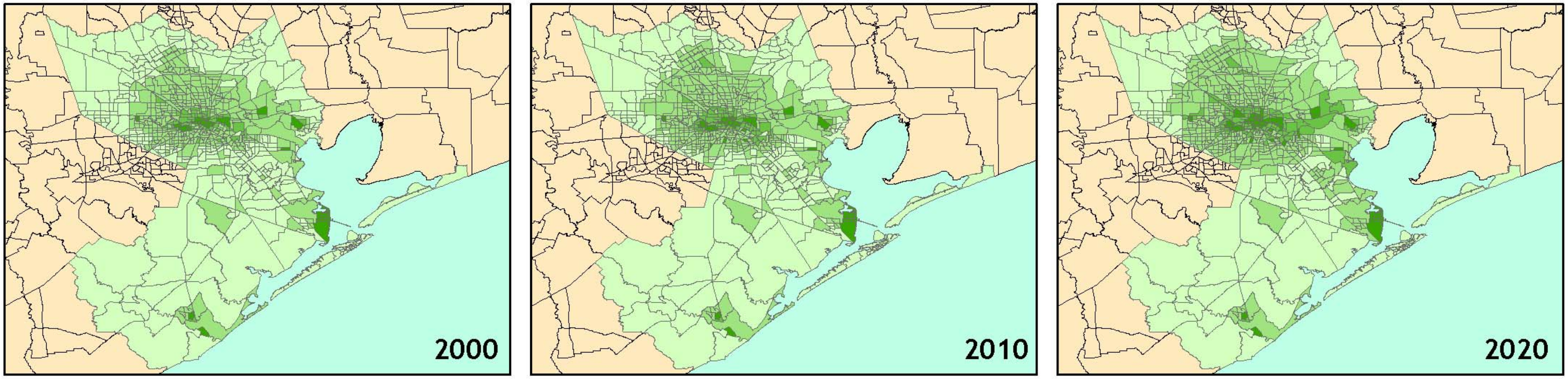


Figure 5 presents maps showing the spatial distribution of benzene reductions across the study area. The top row of maps shows the AERMOD estimates of the reduction in annual average ambient benzene levels due to CAAA programs in (from left to right) 2000, 2010, and 2020. The bottom row shows the same progression using the exposure concentration results from the HAPEM model. On all six maps, the darker shades of green represent greater benzene reductions.

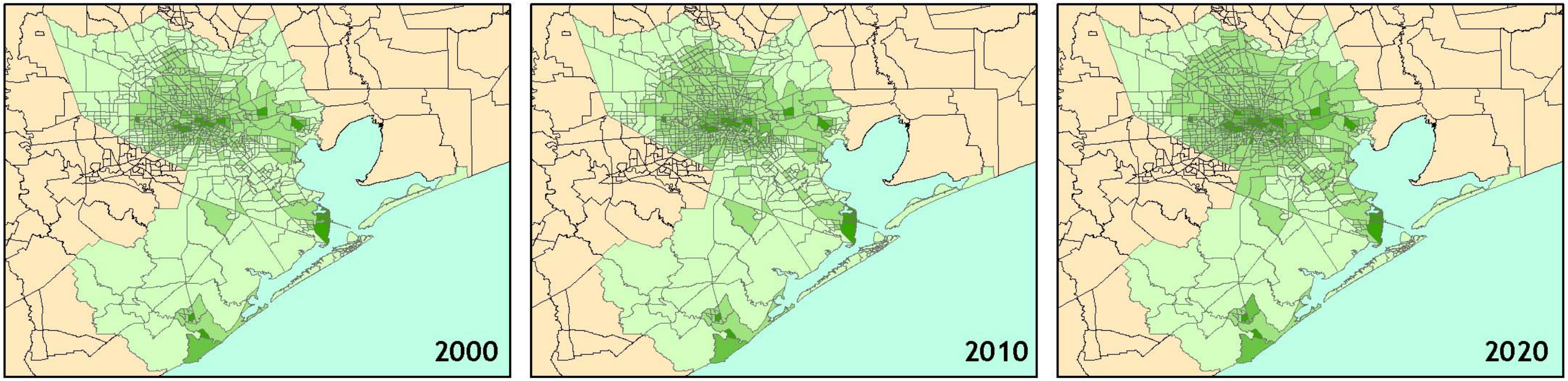
The AERMOD maps show the greatest reductions (in excess of 5 μg/m^3) occur in Harris County in the downtown Houston area, within the rings of the interstate, in the Texas City area of Galveston County where a number of refineries and chemical facilities are located; and in southeastern Brazoria County, which also features major chemical manufacturing and petroleum refining facilities. Mobile source emission controls are a significant contributor to the reductions in Harris County, and thus we can see an increase in the areas experiencing larger reductions in that area, as mobile source controls become more effective over time. The major reductions in Galveston and Brazoria are primarily driven by controls on major point and non-point sources, which tend to be realized earlier in our analysis; thus, we see less change over time in the reductions in these areas. There are some additional reductions gained in the Texas city vicinity, however, most likely due to controls on on-road and non-road sources.

FIGURE 5: ESTIMATED CAAA-RELATED REDUCTIONS IN BENZENE CONCENTRATIONS IN THE HOUSTON METROPOLITAN AREA (*WITHOUT-CAAA* MINUS *WITH-CAAA*) - AERMOD AND HAPEM RESULTS

AERMOD RESULTS



HAPEM RESULTS



Reductions in Concentration 🟩 >2.5 $\mu g/m^3$ 🟩 1.5 to 2.5 $\mu g/m^3$ 🟩 0.5 to 1.5 $\mu g/m^3$ 🟩 <0.5 $\mu g/m^3$

Note: HAPEM results represent the estimated exposure concentration reduction for the median exposed individual in each census tract.

As we compare the maps from top to bottom, we can see the changes in exposure estimates as we process the ambient data through HAPEM to incorporate time-activity patterns of the exposed populations. The exposure changes reflected in the bottom maps represent the change in concentration that we expect would be experienced by the median individual in a given census tract. In general, HAPEM tends to smooth and spread out the AERMOD concentration changes; this reflects both aggregating results to the census tract level and incorporating the impact of commuting and other activities on the concentration experienced by the population in each census tract.

Tables 3 and 4 present mean reductions in annual average benzene from AERMOD and HAPEM, respectively, for each county in each year. In addition, these tables indicate the minimum and maximum reductions estimated for a census block group (AERMOD) or census tract (HAPEM) in that county in that year. To facilitate comparison between the air quality modeling and exposure modeling results, we have calculated population-weighted mean benzene reductions from AERMOD in Table 3. That is, the mean estimates in Table 3 have been adjusted to give more weight to reductions in areas with large populations and less weight to reduction in areas with small populations. The population-weighted mean reductions tend to be around 1 μg/m^3, though the range of reductions can be significant, in several cases exceeding 20 μg/m^3. The results for HAPEM tend to be slightly lower than the AERMOD results. The average ratio of HAPEM to AERMOD concentrations is about 90 percent (see Table 19 in Appendix B), suggesting that much of the population may be commuting from census tracts with higher benzene levels to census tracts with lower levels.

TABLE 3: POPULATION-WEIGHTED MEAN REDUCTION IN AMBIENT ANNUAL AVERAGE BENZENE CONCENTRATION DUE TO CAAA, BY YEAR AND COUNTY

STUDY YEAR	MEAN CHANGE IN BENZENE CONCENTRATION, μg/m^3 (RANGE)		
	BRAZORIA	GALVESTON	HARRIS
2000	1.0 (0.04 - 25)	0.8 (0.04 - 18)	0.8 (-3 - 34)*
2010	1.1 (0.08 - 25)	0.9 (0.05 - 17)	1.0 (-4 - 33)*
2020	1.3 (0.09 - 28)	1.0 (0.06 - 20)	1.2 (-4 - 37)*

* Seven of the 1,911 census block groups in Harris County showed dis-benefits under the *With-CAAA* scenario. Of these, five reported increases of 0.3 μg/m^3 or less. The smallest reductions estimated were between 0.02 and 0.1 μg/m^3.

TABLE 4: HAPEM-ESTIMATED MEAN REDUCTION IN ANNUAL BENZENE EXPOSURE CONCENTRATION DUE TO CAAA, BY YEAR AND COUNTY

STUDY YEAR	MEAN CHANGE IN BENZENE CONCENTRATION $\mu g/m^{3*}$ (RANGE)		
	BRAZORIA	GALVESTON	HARRIS
2000	0.9 (0.07 - 19)	0.7 (0.08 - 14)	0.8 (-1 - 11)**
2010	0.9 (0.1 - 19)	0.7 (0.09 - 14)	0.9 (-1 - 12)**
2020	1.1 (0.1 -21)	0.9 (0.1 - 16)	1.1 (-1 - 14)**

* The HAPEM results in this table represent the exposure change for the median individual in a census tract (i.e., they are neither highly nor minimally exposed in terms of their activities and characteristics). The exposure change is an average change in exposure across all age categories.

* *One of the 649 census tracts in Harris County reported dis-benefits under the *With-CAAA* scenario. The smallest reductions estimated were between 0.07 and 0.1 $\mu g/m^3$.

3.2.1 MODEL TO MONITOR COMPARISONS

The results of the model-to-monitor comparisons are presented in Appendix B. As can be seen in Figures 32 and 33 of that document, many of the AERMOD predicted values fall within a factor of 0.5 to 2 of the monitored values, which is considered good agreement. However, a significant fraction of the *With-CAAA* estimates are less than half of the monitor values, suggesting the model may be underestimating benzene levels.

3.3 HEALTH EFFECTS

This section presents the health effects results and the associated monetary benefits results. We first present the life-table model results for our primary estimate of avoided fatal and non-fatal cases of leukemia (all types) and the monetized value of those cases. We then discuss the results of our assessment of the non-cancer effects of benzene. The next section presents our analysis of CAAA-related individual leukemia risk reductions for individuals that are part of highly exposed populations in the case study area. Finally, we describe the additional life-table model runs we conducted to assess the sensitivity of the model to alternative assumptions.

3.3.1 CANCER

Avoided Cases

Table 5 below presents the results of our primary estimate for avoided fatal and non-fatal cases of leukemia due to CAAA-related changes in ambient benzene levels in the Houston area (including Brazoria, Galveston, and Harris counties). The results are presented for the base year (1990) as well as the three study years (2000, 2010, and 2020). The values in Table 5 represent the annual number of avoided cases in each target year as well as a total number of expected cases avoided from 1990 through 2020. We expect the benefits of the benzene reductions that occur in the study period will continue accruing to the study population beyond the end of the study period. Therefore, we also estimated a total number of cases expected to occur past 2020 that are a result of CAAA-related changes in benzene occurring within the study period. We discuss the derivation of this estimate further in the section entitled "Expected Total Benefits."

TABLE 5: ANNUAL AVOIDED LEUKEMIA CASES (FATAL AND NON-FATAL) BY STUDY YEAR DUE TO CAAA-RELATED BENZENE EXPOSURE CHANGES IN THE HOUSTON AREA

STUDY YEAR	ANNUAL AVOIDED CASES OF LEUKEMIA		
	AVOIDED FATAL CASES	AVOIDED NON-FATAL CASES	TOTAL AVOIDED CASES
1990	0	0	**0**
2000	0.03	0.02	**0.05**
2010	0.09	0.07	**0.2**
2020	0.2	0.1	**0.3**
Cumulative Cases Occurring Within the Study Period	**2**	**2**	**4**
Additional Cumulative Cases Occurring After 2020*	**1**	**1**	**2**
Total Cumulative Cases	**3**	**3**	**6**

*Note: These avoided cases are due to changes in benzene exposure that occurred within the study period.

Our results indicate that by the year 2020, a total of four cases of leukemia would be avoided due to the 1990 CAAA programs in the Houston area, with three of those occurring in Harris County. We estimate two of the four cases to be fatal and two to be non-fatal.[64]

Monetary Valuation

We applied the valuation methods described in Section 2.5.2 to determine the economic value of these avoided leukemia cases. The results of the valuation analysis are presented below in Table 6.

---

[64] The composition of fatal and non-fatal cases is consistent with data from the SEER website for 1988-2004, which indicates that ten year-survival rates for leukemia are approximately 40 percent (http://seer.cancer.gov/).

TABLE 6: ANNUAL MONETARY BENEFITS BY STUDY YEAR DUE TO CAAA-RELATED CHANGES IN BENZENE EXPOSURE IN THE HOUSTON AREA

STUDY YEAR	TOTAL BENEFITS (1990 NPV, MILLIONS OF 2006$, 5% DR)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
1990	$0	$0	**$0**
2000	$0.12	$0.01 – 0.06	**$0.13 - 0.18**
2010	$0.27	$0.01 – 0.13	**$0.28 - 0.40**
2020	$0.31	$0.01 – 0.15	**$0.32 - 0.46**
Cumulative Cases Occurring Within the Study Period	**$6.7**	**$0.32 - 3.3**	**$7.0 - 10**
Additional Cumulative Cases Occurring After 2020*	**$1.6**	**$0.08 - 0.8**	**$1.7 - 2.4**
Total Cumulative Cases	**$8.3**	**$0.40 - 4.1**	**$8.7 - 12**

*Note: These avoided cases are due to changes in benzene exposure that occurred within the study period, but occurred after 2020 due to lagging effects of these changes on leukemia risks.

The values in Table 6 represent the annual net present value estimate (discounted to 1990) of the benefits of the CAAA-related benzene controls in Houston in each target year.[65] In addition, we calculated the net present value of benefits over the entire study period and the additional benefits of cases occurring after 2020. Our primary estimate of total benefits due to CAAA-related reductions in benzene are $8.7 - 12 million (in 2006$), $8.3 million of which are due to fatal cases of leukemia, and $0.4 – 4.1 million of which are due to non-fatal cases. Our primary estimate incorporates a discount rate of 5 percent to account for the effect of cessation lag on the distribution of benefits over time.

Expected Total Benefits

The life-table model we applied in this analysis was designed to calculate the change in the number of cases of leukemia likely to be observed in a given year, as a function of a population's current and past exposures. Because of the way we model the lag between exposure reduction and benefits (see Section 2.4.2), the exposure change in the year being modeled contributes little to the observed risk reduction in that year; most of its effects will be realized over the next several years. Similarly, the exposure changes in the years preceding the year being modeled will continue to produce benefits in future years, to a lesser degree over time. As a result, a portion of the benefits that result from exposure changes that occur in the 1990 to 2020 study period will not be observed until after 2020.

---

[65] Net present value (NPV) calculations facilitate comparison of costs or benefits that may occur at different points in the future by expressing them in terms of their value in a common reference year, using the economic principle of discounting. For example, the value of X dollars received N years from today would be $X/(1 + i)^N$, where i represents the discount rate, a measure of the time value of money. In this case study, we discount the value of all future health benefits back to the first year of the analysis, 1990, and sum them to produce our NPV estimates.

To address this model limitation, we estimated the relative magnitude of the benefits that we expected would occur after the end of the study period (i.e., past the year 2020), assuming that the latency period assumed in our primary estimate is correct. In order to generate an estimate of the size of these benefits, we ran the model using a truncated exposure data set that "turned off" the effect of the CAAA after 2010 (i.e., it assumed no difference in exposure between the *With*- and *Without-CAAA* scenarios after the year 2010) and observed how the benefits decreased following 2010. We found that annual avoided cases peaked in the year 2010 and then decreased to 90 percent of the 2010 level for the first five years (2011-2015) and to 50 percent of the 2010 level for the next 5 years (2016-2020). (Although we did not model past 2020, we believe the benefits after 10 years will likely be minimal, given the exposure weights we used in the model.) We believe the decay in benefits observed in this example 2010 cutoff run represent a reasonable approximation of the results that would be observed after 2020.

We applied the ratios from the 2010 cutoff run to the 2020 estimates of annual avoided cases and calculated estimates of cumulative avoided cases for 2025 and 2030. We estimated less than one additional fatal case and less than one additional non-fatal case avoided in the first five years after the study period. By the year 2030, we estimated another partial fatal and another partial non-fatal case would be avoided, making the cumulative total cases avoided through 2030 due to benzene concentration changes between 1990 and 2020 to be roughly six.

3.3.2 NON-CANCER

As described in Section 2.4.5, in order to assess non-cancer health benefits, we planned to report the difference between the *With-CAAA* and *Without-CAAA* scenarios in the number of individuals experiencing benzene concentrations above the chronic RfC published in EPA’s Integrated Risk Information System (IRIS) database. Therefore, we compared the chronic RfC value reported on IRIS (0.03 mg/m^3) with the ambient benzene concentrations from HAPEM6 for each tract under both the *With*- and *Without-CAAA* scenarios. We then calculated the total census population across all of the tracts with benzene concentrations exceeding the RfC under each scenario. We found no individuals exposed to benzene at concentrations exceeding the RfC in either the *With*- or *Without-CAAA* scenarios.

3.3.3 HIGHLY-EXPOSED POPULATIONS

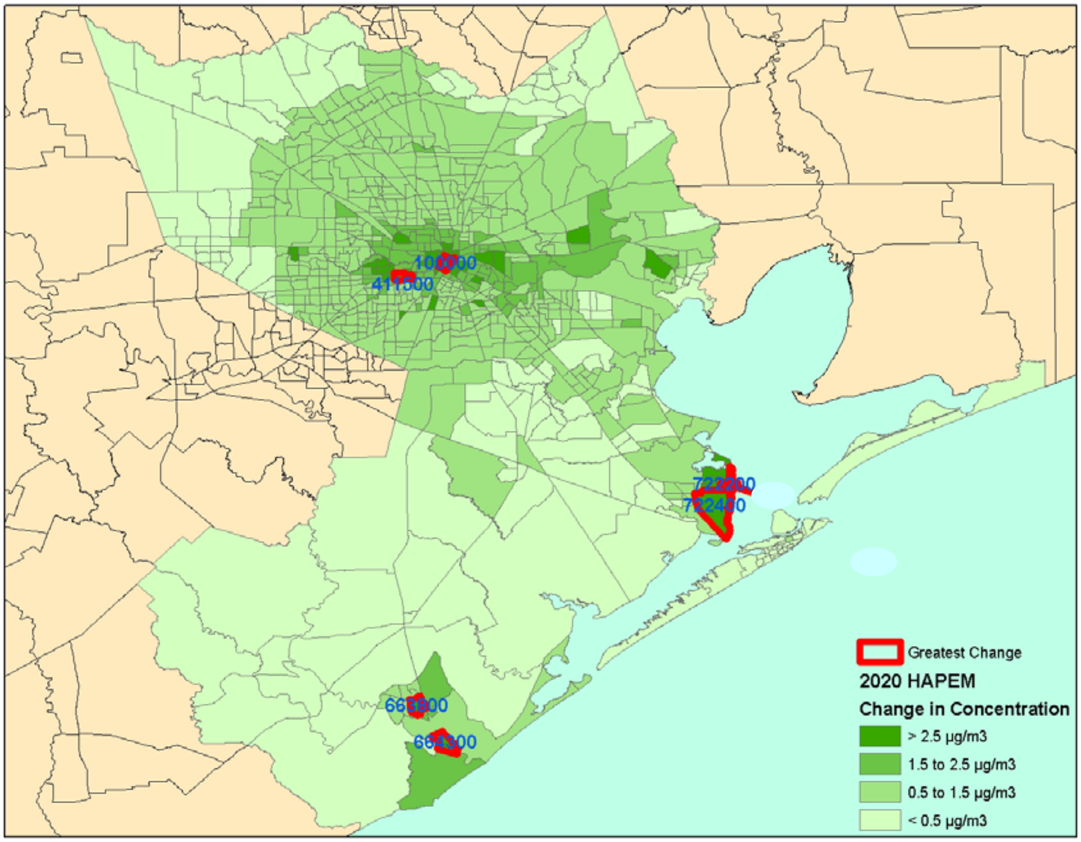
We evaluated risks to three different sets of highly exposed populations: residents living in census tracts with high benzene concentrations, residents living near roadways, and residents living in homes with attached garages.

Residents Living in Census Tracts With High Exposure

As described in Section 2.4.4, we estimated CAAA-related reductions in the lifetime risk of leukemia for individuals living in census tracts with the highest levels of benzene. Figure 6 shows a map that highlights these census tracts. Table 7 below presents the individual lifetime risk of leukemia for a person born in 2020 under both the *With*- and

Without-CAAA scenarios in the two tracts in each county with the highest exposure concentrations. In addition, we report the population of these tracts, who would experience these levels of risk or higher. Risks under the *Without-CAAA* scenario are significantly higher compared to those under the *With-CAAA* scenario. For example, some risks in Brazoria County drop from an increased lifetime leukemia risk of 2 in ten thousand (i.e., 2×10^{-4}) to 3 in a million (3×10^{-6}) as a result of the CAAA, a 98 percent reduction. In four of the six tracts in Table 7, individual lifetime leukemia risks among the highly exposed are reduced by at least 80 percent; the risks in all six counties are reduced by at least 72 percent.[66] For comparison, the estimated average lifetime leukemia risk reduction across the 3-county study area for an individual born in 2020 is 65 percent.

FIGURE 6: CENSUS TRACTS IN THE HOUSTON STUDY AREA WITH THE GREATEST BENZENE EXPOSURE CONCENTRATION CHANGES BETWEEN THE *WITH-* AND *WITHOUT-CAAA* SCENARIOS IN 2020



---

[66] Risks were calculated using the 7.8×10^{-6} per µg/m^3 benzene inhalation unit risk (IUR) from the range of IURs reported on IRIS.

TABLE 7: CAAA-RELATED LEUKEMIA RISK REDUCTIONS IN 2020 IN THE HOUSTON AREA FOR INDIVIDUALS LIVING IN CENSUS TRACTS WITH HIGH AMBIENT BENZENE CONCENTRATIONS

COUNTY	CENSUS TRACT	MEDIAN WITHOUT-CAAA RISK	MEDIAN WITH-CAAA RISK	PERCENT REDUCTION IN RISK	POPULATION OF CENSUS TRACT
Brazoria	6643	2×10^{-4}	3×10^{-6}	98	5,452
Brazoria	6638	3×10^{-5}	6×10^{-6}	77	4,470
Galveston	7222	1×10^{-4}	7×10^{-6}	95	3,487
Galveston	7224	5×10^{-5}	8×10^{-6}	82	1,108
Harris	1000	1×10^{-4}	1×10^{-5}	92	6,678
Harris	2523	3×10^{-5}	7×10^{-6}	72	12,686

Note: These risk values were calculated using the 7.8 × 10-6 per μg/m^3 benzene inhalation unit risk (IUR) from the range of IURs reported on IRIS.

Residents Living Near Roadways

Figure 7 displays boxplots of the results of our 2002 HAPEM runs with and without the near-roadway algorithms. We present results for both the median (50th percentile) and highly exposed (90th percentile) individual.

The boxplots on the left show little change in benzene reductions for the median exposed individual after applying the near-roadway algorithms. Our primary benefit estimates, which are based on the median exposure results, therefore reflect minimal impact of the near roadway adjustment. This is not surprising, because it is unlikely that half of the study population would live within 75 or 200 meters of a major roadway. However, on the right side of Figure 7, we do see an increase in benzene reductions for highly exposed individuals after applying the near-road algorithms. The entire distribution of benzene reductions for the highly exposed group shifts upward, and the median reduction in benzene exposure for this group is about 20 percent larger than the run with the near-roadway algorithm turned off. Thus, overall for the highly exposed group, we observe a moderate impact of incorporating near-roadway effects on benefits. An analysis of the ten census tracts with the highest on-road-related benzene concentrations in 2020 under the *Without-CAAA* scenario (and total population greater than 100) shows more significant impacts in individual locations, with the exposure reduction in one tract in Harris County nearly doubling. On average, the exposure (and hence, risk) reductions in these ten tracts for highly exposed individuals are one and a half times larger when the near-roadway effect is taken into account.

FIGURE 7: BOXPLOTS OF CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA IN 2020 - IMPACT OF INCORPORATING NEAR-ROADWAY EFFECTS

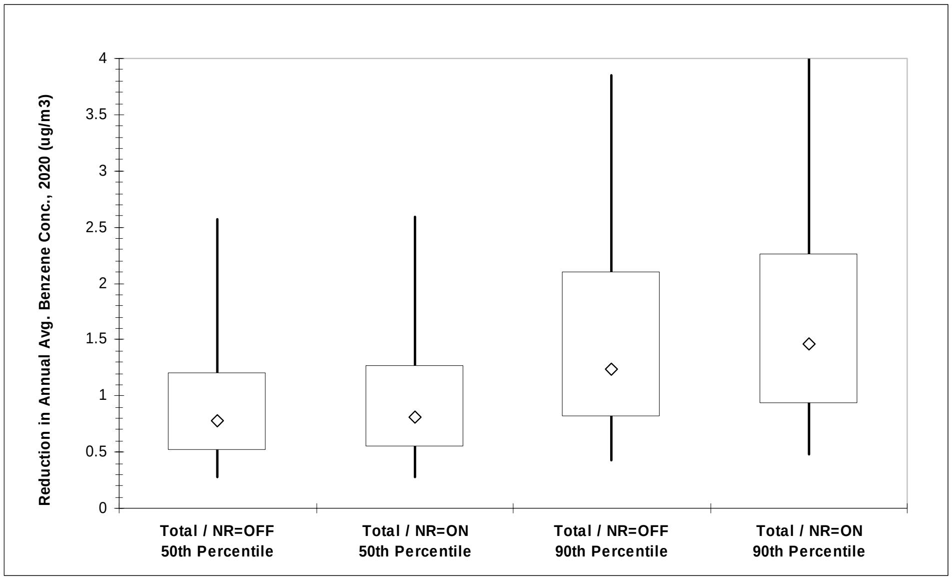


TABLE 8: CAAA-RELATED BENZENE REDUCTIONS IN 2020 INDIVIDUALS LIVING IN CENSUS TRACTS WITH HIGH AMBIENT BENZENE CONCENTRATIONS DUE TO ON-ROAD SOURCES - EFFECT OF HAPEM NEAR-ROADWAY ALGORITHM

COUNTY	CENSUS TRACT	BENZENE REDUCTION NEAR-ROADWAY OFF ($\mu g/m^3$)	BENZENE REDUCTION NEAR-ROADWAY ON ($\mu g/m^3$)	PERCENT CHANGE IN BENZENE DUE TO NEAR ROADWAY EFFECT	POTENTIALLY AFFECTED POPULATION[1]
Harris	321500	1.5	2.6	69	226
Harris	540200	1.3	2.5	89	247
Harris	310700	2.3	3.8	65	457
Harris	541900	2.0	2.5	25	436
Harris	431200	2.4	3.5	44	694
Harris	412100	1.6	2.5	60	98
Harris	450300	2.2	3.1	43	712
Harris	311900	2.0	2.8	42	278
Harris	431900	3.0	3.5	15	206
Harris	410900	2.7	3.3	21	282

[1] Because these values were calculated using 90th percentile exposure concentrations, we assumed that 10 percent of the population in the tracts may be associated with these changes in benzene exposure or higher.

Residents With Attached Garages

We estimated that total emissions in attached garages in the Houston area would be reduced by almost 90 percent. If the average exposure estimate attributable to attached garages (1.2 $\mu g/m^3$; see USEPA, 2007d) were reduced by this amount, the expected reduction in exposures due to reductions of in-garage emissions would be 1.1 $\mu g/m^3$. We found that this would correspond to an additional estimate of annual avoided cases of leukemia in the Houston area in 2020 that is roughly similar in magnitude to our main benefits estimate. Therefore, these results suggest that adding attached garage-related benefits to our primary estimate could result in an approximate doubling of our primary estimate.[67]

3.3.4 SENSITIVITY ANALYSES

We performed five sensitivity analyses to estimate the range of uncertainty surrounding our primary estimate and to determine how sensitive the health risk model is to various data inputs and assumptions.[68] We first assessed the impact of statistical uncertainty surrounding our primary estimate by running the model with the upper and lower 95 percent confidence limits of the dose-response slope factor from Crump (1994). We then tested the sensitivity of the model to the underlying epidemiological data by substituting the dose-response slope factor used in our primary estimate with that from another major cohort study linking benzene and leukemia. Next, we explored the sensitivity of the model to the health endpoint selected by looking at the differences between incidence rates for all leukemia versus AML. We next substituted a dose-response slope factor derived using different exposure estimates from the same cohort used in our primary estimate, the Pliofilm Cohort. We also ran the model with two alternate lags, a zero-year lag and a five-year lag.

We also explored the range of uncertainty surrounding assumptions made during the valuation of the health effects results. We performed a sensitivity analysis on our primary valuation estimate by altering the discount rate applied. We also substituted alternate values for the VSL used to value fatal cases of leukemia. Finally, we assumed that all of the leukemia cases due to benzene exposure were fatal to get an upper bound benefits estimate.

Statistical Uncertainty

Our primary estimate of avoided cases of leukemia relied on a mean dose-response slope factor from the Crump (1994) paper. To assess the impact of statistical uncertainty on this estimate, we ran the life-table model with both the upper and lower 95 percent

---

[67] Homes with attached garages may also experience significant short-term spikes in benzene concentrations in the house following cold start or hot soak events (Graham and Noseworthy, 2004). CAAA controls would also be expected to reduce these acute benzene exposures to individuals living in these homes; however estimation of these benefits is beyond the scope of this analysis.

[68] We did not perform a Monte Carlo analysis as part of the sensitivity analysis due to the large amount of data involved and time and resource limitations

confidence bounds (UCL and LCL) around the mean dose-response slope factor, as reported in Crump (1994).

These additional runs indicate that based solely on the statistical uncertainty in the selected dose-response function from the Pliofilm cohort, total cumulative avoided cases of leukemia occurring within the study period could range from a lower bound of 0.8 to an upper bound of seven.

Chinese Worker

Our primary estimate of avoided leukemia cases from the life-table model relied on dose-response slope factors for the relationship between benzene and leukemia from the Pliofilm Cohort, as these are the data currently supported by EPA in the benzene IRIS profile to calculate potency estimates. For our sensitivity analysis, we used a dose-response slope factor from another large, well-studied occupational cohort, the Chinese Worker Cohort. The strengths of this cohort study include a large number of leukemia cases and workers who were exposed to benzene levels similar to ambient levels.

Because the studies examining the Chinese Worker Cohort did not derive dose-response slope factors, we used dose-response slope factors derived by the California Environmental Protection Agency (CalEPA) as part of an analysis to calculate a Public Health Goal for benzene (CalEPA, 2001).[69] We also applied the same lag to our exposure data as was assumed in the Chinese Worker Cohort (1.5 years). The life-table model run with this alternate dose-response slope factor and 1.5-year lag estimated that a total of seven cases of leukemia would be avoided between 1990 and 2020 due to the CAAA.

AML

Our primary estimate was based on a dose-response slope factor derived with all leukemia as the health endpoint. To test the sensitivity of this assumption, we first compared rates for all leukemia to those for AML, the leukemia subtype with the most data supporting its link with benzene, to estimate the proportion of leukemia cases that were AML. We compared national-level age-specific AML incidence rates to national age-specific all leukemia incidence rates.[70] We found that the age-specific all leukemia incidence rates were on average four times higher than the AML rates and ranged from two times higher (for the 25-29 age group) to nine times higher (for the 5-9 age group).

---

[69] The CalEPA dose-response slope factors were derived by applying Poisson regression to relative risks presented in Hayes et al. (1997) and were based on an analysis of a subset of the Chinese Worker Cohort (representing approximately 76 percent of the total person-years at risk) for which exposures remained relatively constant over their work experience, making their exposure assignments less uncertain (CalEPA, 2001). We selected the dose-response slope factor that assumed a linear dose-response function for extrapolation to low doses, as the data was not inconsistent with a linear model. In addition, EPA's *Guidelines for Carcinogen Risk Assessment* (USEPA, 2005b) state that linear extrapolation should be used when the mode of action is uncertain, which is the case for benzene. Given the low concentrations that are likely to be experienced in our case study, a linear approximation may be a reasonable fit, even if the overall dose-response function in supralinear, provided the data from which the extrapolation is being made are not in the plateau region of the curve. In this case, the linear slope might be too shallow, underestimating the true dose-response relationship at low doses. To address this, the CalEPA analysis excluded data points expected to be in the plateau region of the curve.

[70] The source of the AML and all leukemia national incidence rates was the Center for Disease Control (CDC) WONDER online database. See http://wonder.cdc.gov/.

To estimate avoided cases of AML, we multiplied the leukemia incidence rates by ¼ and ran the model using the dose-response slope factor derived using AML as the health endpoint in Crump (1994). We found that the incidence results for AML were 70 percent of the all leukemia results. Therefore, we would expect a total of three avoided cases of AML (fatal and non-fatal) between 1990 and 2020 due to CAAA-related changes in benzene exposure. Ten-year survival rates for AML are approximately 20 percent. These data can be used as an approximation for how many cases are expected to be fatal and non-fatal. Therefore, we would expect that of the three avoided cases of AML, approximately two would be fatal and one would be non-fatal.

Alternate Exposure Matrix

Exposure assessment for the Pliofilm Cohort has been investigated by three separate research groups, Rinsky et al. (1981 & 1987), Crump and Allen (1984), and Paustenbach et al. (1992), yielding a variety of results. The different exposure assessment results of these three analyses can be attributed to various assumptions made by the investigators in relation to exposure of the workers, such as exposure concentrations experienced before sufficient monitoring data was available. Paustenbach et al. estimates are the highest, followed by Crump and Allen, and then Rinsky et al. Accordingly, the Paustenbach et al. estimates yield lower relative risks than the other two exposure estimates.[71] These dose-response slope factors assumed the same health endpoint (all leukemia) and lag (weighted) as the primary estimate. We found much lower health benefits using the Paustenbach exposure estimates, with only two cases of leukemia avoided between 1990 and 2020.

Alternate Lag

Our primary estimate relied on a "weighting" scheme to calculate a cumulative exposure value, with the peak weight being applied 5.3 years prior to the current year as an estimate of the latency period for leukemia. We also ran the model using alternative risk models that assumed a different lag structure. Because the lag structure is an integral part of how the risk coefficient is estimated in the benzene epidemiological analyses, different lag structures also imply different risk coefficients. We applied two models from Crump (1994), one derived assuming that all previous exposures were weighted equally (with no lag) and the other derived assuming all previous exposure were weighted equally with the exception of the most recent five years, which were weighted with zero. In addition to applying the alternative dose-response slope factors from these risk models, we also applied the corresponding exposure weights from each model to the exposure values from

---

[71] The estimates by Paustenbach et al. (1992) have been criticized for being based upon worst-case assumptions for the exposure scenarios that existed during the early years of the cohort (Utterback and Rinsky, 1995). In fact, critics have noted that prolonged exposure to the high levels of benzene estimated by Paustenbach et al. would have resulted in much higher prevalence of benzene poisoning than was actually seen in the cohort. Nevertheless, we performed a sensitivity analysis using dose-response slope factors from the Crump (1994) analysis derived using the Paustenbach exposure matrix to test the model's sensitivity to this input.

HAPEM6.[72] The dose-response slope factors associated with the zero- and five-year lags are lower than the dose-response slope factor used for the primary estimate (0.017 versus 0.84), in part because the weighted exposure values for these lag models are considerably higher than for our main model. The effect of the lower coefficient counteracts the effect of the shorter lags, and apparently has a greater impact; the results we found for these alternate lags were lower than the primary estimate. The zero-year lag model run yielded an estimate of two avoided cases between 1990 and 2020 and the five-year lag yielded an estimate of one case.

Discount Rate

We also estimated total monetary benefits using alternative discount rates of 0, 3, and 7 percent, as described in Section 2.5.2. The results of this sensitivity analysis are presented in Table 9 and range from $4.9 – 7.1 million for the high discount rate to $19 – 27 million when no discount rate is applied.

TABLE 9: TOTAL BENEFITS DUE TO CAA-RELATED CHANGES IN BENZENE OCCURING WITHIN THE STUDY PERIOD, CALCULATED WITH ALTERNATIVE DISCOUNT RATES

DISCOUNT PERCENTAGE	TOTAL BENEFITS (1990 NPV, MILLIONS OF2006$)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
Primary Estimate (5%)	**$6.7**	**$0.3 - 3.3**	**$7.0 - 10**
No Discounting	$18	$0.9 – 9.0	$19 – 27
Low Discount Rate (3%)	$9.8	$0.5 – 4.9	$10 – 15
High Discount Rate (7%)	$4.7	$0.2 – 2.3	$4.9 – 7.1

VSL

We selected a VSL of $7.4 million in 1990 (2006$) for our primary estimate, from a 2003 meta-analysis of wage-risk studies by Viscusi and Aldy (Model 5 from Table 8). To explore the sensitivity of the results to this assumption, we calculated the economic benefits using the following alternative VSL estimates:

- An alternative estimate from Viscusi and Aldy (2003) (Model 2 from Table 8) that assumes a log-normal distribution with a mean of $5.8 million (in 2000$);
- The estimate used in the recent PM NAAQS RIA assuming a normal distribution with a mean of $5.5 million (in 2000$); and
- An estimate used by EPA in past benefits analysis assuming a Weibull distribution based on 26 studies, with a mean of $4.8 million (in 1990$).

---

[72] For example, for the five-year lag, we applied a weight of 0 to the most recent five years of exposure and a weight of 1 to all other past exposures within the study period.

The total benefits estimated using these alternative VSL estimates, converted to 2006 dollars, are displayed in Table 10 below.

TABLE 10: TOTAL BENEFITS DUE TO CAA-RELATED CHANGES IN BENZENE OCCURING WITHIN THE STUDY PERIOD, CALCULATED WITH ALTERNATIVE VSL ESTIMATES

VSL	TOTAL BENEFITS (1990 NPV, MILLIONS OF2006$)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
Primary Estimate (Viscusi and Aldy, 2003, Model 5)	**$6.7**	**$0.3 - 3.3**	**$7.0 - 10**
Viscusi and Aldy, 2003, Model 2	$6.2	$0.3 – 3.1	$6.5 – 9.3
Normal Distribution	$5.9	$0.3 – 2.9	$6.2 – 8.8
Weibull Distribution	$6.7	$0.3 – 3.3	$7.0 – 10

Fatality Rate

In our primary estimate, we assumed that the difference between running the model with incidence data and mortality data constituted the number of leukemia cases that would be non-fatal. We found that of the four avoided cases of leukemia that would occur between 1990 and 2020, two would be fatal and two would be non-fatal (i.e., a 50 percent fatality rate). Although ten-year survival data for 1988-2004 presented on the SEER website supports this (the data indicate a 60 percent fatality rate within ten years), it is possible that those that survive ten years could come out of remission and eventually die of leukemia. In order to test the sensitivity of the results to this assumption, we calculated an alternate estimate of the monetary benefits assuming that all cases were fatal. We found that the total monetary benefits would increase to $13 million (in 2006$) using a five percent discount rate.

Summary

Table 11 displays annual avoided cases (fatal and non-fatal) of leukemia by study year and total cumulative cases occurring within the study period for the primary estimate as well as estimates for the sensitivity analyses. Total avoided cases between 1990 and 2020 for the primary estimate is four and the sensitivity analyses range between one and seven. Figure 8 presents the annual avoided cases of leukemia between 1990 and 2020 for the primary case as well as five of the sensitivity analyses in graphical form.

We also assessed the economic benefits associated with the avoided cases of leukemia for the sensitivity analyses. Table 12 below presents the total monetary benefits (for both fatal and non-fatal cases of leukemia) for the primary case as well as the sensitivity analyses.

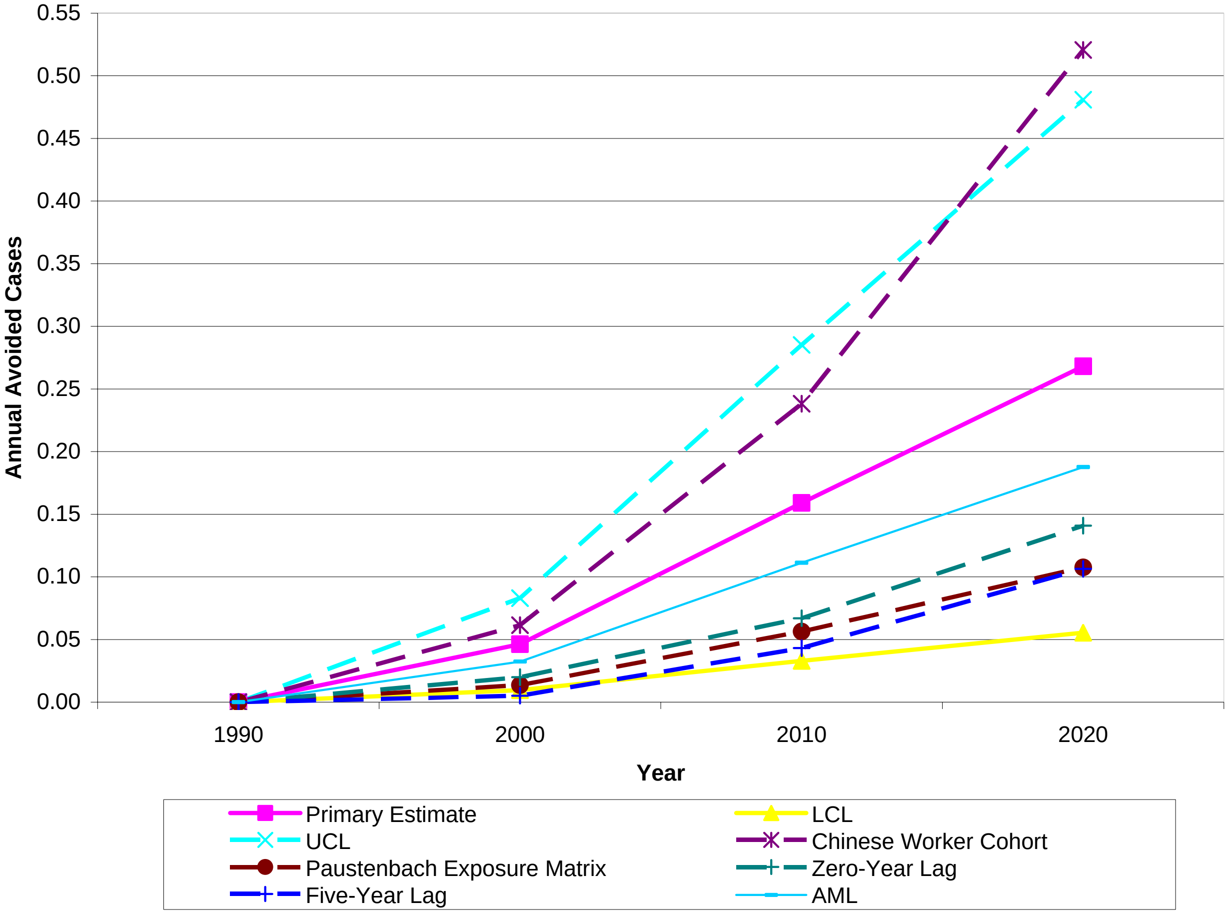
TABLE 11: TOTAL AVOIDED CASES OF LEUKEMIA DUE TO CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA – PRIMARY ESTIMATE AND SENSITIVITY ANALYSES RESULTS

YEAR	PRIMARY ESTIMATE	LCL	UCL	CHINESE WORKER COHORT	PAUSTENBACH EXPOSURE MATRIX	ZERO-YEAR LAG	FIVE-YEAR LAG	AML
1990	**0**	0	0	0	0	0	0	0
2000	**0.05**	0.01	0.08	0.06	0.01	0.02	0.01	0.03
2010	**0.2**	0.03	0.3	0.2	0.06	0.07	0.04	0.1
2020	**0.3**	0.06	0.5	0.5	0.1	0.1	0.1	0.2
Total Cumulative Cases	**4**	0.8	7	7	2	2	1	3

TABLE 12: TOTAL MONETARY BENEFITS OF CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA – PRIMARY ESTIMATE AND SENSITIVITY ANALYSES RESULTS (IN MILLIONS OF 2006$)

PRIMARY ESTIMATE	LCL	UCL	CHINESE WORKER COHORT	PAUSTENBACH EXPOSURE MATRIX	ZERO-YEAR LAG	FIVE-YEAR LAG	ALL FATAL
$7.0 - 10	$1.5 – 2.0	$13 - 18	$11 – 16	$2.5 – 3.6	$3.2 – 4.6	$2.0 – 2.8	$13

FIGURE 8: ANNUAL AVOIDED CASES OF LEUKEMIA DUE TO CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA - PRIMARY ESTIMATE AND SENSITIVITY ANALYSES RESULTS



Note: We have linearly interpolated between the avoided leukemia estimates for each target year; however, the true shape of the curve between each of these points is uncertain.

CHAPTER 4 | DISCUSSION AND CONCLUSIONS

This chapter discusses the key findings of this case study and the uncertainties associated with its results. It also presents an assessment of the strengths and limitations of the modeling approach used in this analysis and its implications for potential future assessment of the benefits of HAP controls.

4.1 KEY FINDINGS

This case study demonstrates that the 1990 CAAA controls on benzene emissions are expected to result in reductions in the incidence of leukemia in the greater Houston area over the period 1990 to 2020. Key findings include:

- CAAA programs are expected to reduce benzene emissions across all source categories in the study area by thousands of tons per year, with the largest reductions in the point and non-point source category, followed by on-road and non-road sources;
- The largest reductions in benzene exposures are expected to occur in downtown Houston and the surrounding area, and in two areas with significant point sources: the Texas City area of Galveston County and southeastern Brazoria county;
- Reductions in benzene levels are expected to continue, and hence benefits are expected to increase in the latter decades of the study period, as engine and other capital stock turns over and the impact of CAAA controls on on-road and non-road mobile sources in the area increases;
- Primary benefit estimates indicate four fewer cases of leukemia would occur in the three-county area in the study period, two of which we expect would have been fatal. We also expect benefits from the benzene changes that occur between 1990 and 2020 will continue accruing through at least 2030, potentially avoiding another two leukemia cases between 2020 and 2030. We estimate the net present value (NPV) in 1990 of the two fatal and two non-fatal leukemia cases avoided to be between $7.0 – 10 million in 2006 dollars, using a five percent discount rate.
- 1990 CAAA controls on benzene are expected to significantly reduce individual leukemia risk levels for those living in census tracts with the highest estimated benzene levels by one to two orders of magnitude. For example, median risks in Brazoria County drop from an increased lifetime leukemia risk of 2 in ten thousand (i.e., 2×10^{-4}) to 3 in a million (3×10^{-6}). In four of the six census tracts with the highest risks, individual lifetime leukemia risks are reduced by at least 80 percent.

- Additional health benefits may accrue to individuals living in homes with attached garages. Back-of-the-envelope estimates of the benefits of CAAA-related benzene reductions in the garages of these homes suggest these benefits may be similar in magnitude to our primary estimate. Therefore, these results suggest that adding attached garage-related benefits to our primary estimate could result in an approximate doubling of our primary estimate.

Although the actual benefit results appear modest, we note that leukemia is a rare disease with a low baseline rate among the population - for people under 50, the baseline risk in the study area was generally less than 5 in 100,000. Therefore, even significant percentage reductions in the baseline leukemia mortality rate may translate to relatively small numbers of avoided cases. We also note that the cases avoided are associated with only three U.S. counties containing just over one percent of the total U.S. population. We would expect significantly higher numbers of leukemia cases avoided when looking nationally at benzene reductions.

4.2 UNCERTAINTIES AND DATA GAPS

The results of this case study reflect limitations in available data and resources for conducting this analysis, as well as in the models and assumptions inherent in our analysis. Where feasible, we have conducted quantitative analysis to estimate potential impacts of these uncertainties; in other cases, we discuss qualitatively the source of uncertainty and our best estimate of the direction and size of its potential impact. We believe that overall, the uncertainties in our analysis are likely to cause our results to be underestimated.

We reach this conclusion for several reasons. First, the apparent systematic underestimation of benzene levels throughout the study area, due to upstream uncertainties in emissions and air quality modeling, constitutes one of the most major biases affecting our results. Further down in the analytical chain, additional factors contributing to a downward bias include the exclusion of a number of potential benzene-related health endpoints that we were unable to quantify for this case study; the exclusion of a ME for attached garages in the exposure modeling step, and uncertainties in the appropriate model for the C-R relationship between benzene exposure and leukemia. We describe the potential uncertainties of the study in greater detail below.

4.2.1 ESTIMATION OF BENZENE LEVELS

We believe that modeled benzene levels in this case study on average underpredict true ambient levels. Comparison of modeled ambient benzene levels from the *With-CAAA* AERMOD run for the year 2000 with observed monitor results in Appendix B shows a significant fraction of results are less than half of the observed values. These low results may be due to uncertainty or bias in emissions estimates, in air quality modeling, or a combination of the two. We discuss below the uncertainties we believe are likely to have a more significant impact on results. For more detail on uncertainties in emissions and air quality and exposure modeling, please consult Appendices A and B, respectively.

On the emissions side, air quality studies in Southeast Texas (TexAQS I and II) have led researchers to conclude that there is a high level of uncertainty in the HG area point source VOC and air toxic contaminant emission estimates, especially for petrochemical facilities.[73] As discussed in Appendix A, TexAQS II in 2006 confirmed that inventories based on standard EPA emission factors significantly underestimated VOC emissions from petrochemical facilities; it found that ethane emissions from petrochemical facilities were underestimated in the 2004 TCEQ point source database by one or two orders of magnitude.[74] Airborne measurements of VOCs, including benzene, from TexAQS II further support this hypothesis – measurements as high as 50 ppb benzene were detected in the Houston Ship Channel area where many petrochemical facilities are located, concentrations which are not consistent with emissions reported in the area's point source inventories. Given the significant contributions of point source sector regulations to the overall benzene reductions observed in the case study, this has the potential to be a major source of bias. However, the true impact of this downward bias depends on how well the missing emissions would be controlled by CAAA-related regulations, such as the highly reactive volatile organic compound (HRVOC) rules initiated in the study area since 2000. If these rules are effective in reducing benzene emissions from fugitive emission sources, we would underestimate the benefits of the CAAA. In its review of this case study, the SAB (USEPA, 2008b) strongly emphasized the need for additional investigation into missing or underestimated HAP emissions categories for point sources in Houston or other major metropolitan areas with significant point source HAP emissions to help reduce this bias in future assessments.

Another source of downward bias in emissions is the omission from the *With*-CAAA scenario of industrial leak detection and repair reductions that are part of the Texas SIP for ozone. Because these programs have been adopted in order to reduce fugitive VOC emissions that have not been captured in the VOC emission inventory for the study area, their emission benefit is difficult to model. Emissions for source categories affected by LDAR rules are likely underestimated. Additional research into the potential magnitude of these emissions reductions would benefit future studies. Of the other categories of benzene emissions controls that were not included in our analysis, the most significant is likely the set of controls associated with the MSAT program, which was established after the *With-CAAA* scenario was fixed.[75]

Reduction of benzene emissions in the mobile source category constitutes another significant contribution to CAAA benefits. Comparison of the base on-road inventory we used (the 2002 NEI on-road inventory) with the on-road emissions from the 1999 NEI showed significantly lower emissions in our inventory (about 760 tpy) than those from

---

[73] See, for example, Ryerson et al., 2003, Kleinman et al., 2002, and Allen and Durrenberger, 2003.

[74] For more information on TexAQS II, see http://www.tceq.state.tx.us/nav/eq/texaqsII.html.

[75] Other control categories not addressed by this case study include portable fuel containers, which may contribute to attached garage-related exposures, and new evaporative emissions categories such as tank and hose permeation included in the most recent NONROAD model (NONROAD2005). We also note that cold temperature start emissions for Tier 1 and later vehicles are underestimated by MOBILE6 (USEPA, 2007a); however this is not likely to be a major factor in the warmer Houston climate.

the 1999 NEI (about 1,940 tpy). Further investigation of this discrepancy identified three major contributors: use of local input data for the vehicle registration distribution for the 2002 inventory, revised 2002 summer fuel benzene levels, and reductions from control programs between 1999 and 2002 (Cook, 2007). This comparison illustrates that our results are highly sensitive to fleet distribution and fuel benzene content assumptions. While it is possible this may also contribute to the downward bias, we believe the selection of the 2002 inventory to generate on-road emissions was reasonable given its use of local, rather than national, registration data and its use of more up-to-date data and assumptions.

On the air quality modeling side, model-to-monitor comparisons suggest our AERMOD runs may have underestimated ambient benzene concentrations in 2000, as more than a quarter of the estimates are less than half the corresponding monitor values. If the air quality modeling systematically underestimated concentrations for both scenarios, it is possible that the difference between the two scenarios may also be underestimated, biasing our benefits estimates downward. If the size of the modeling error is approximately constant, the error would be subtracted out when we calculate the difference between the two scenarios and would not affect our results. If however, the error is proportional to the magnitude of the concentration modeled, then the error could result in an underestimate the difference between the scenarios.[76]

As noted by the SAB during its review of this case study (USEPA, 2008b), the lack of modeling of benzene concentrations during calm periods (“calms”), when high exposures are expected to occur, is likely a contributor to the observed downward bias. AERMOD by design is unable to estimate concentrations during calms (i.e., zero wind speed), and there is some uncertainty related to how well AERMOD performs when one substitutes a very low wind speed (e.g., less than 1m/s) for a calm. EPA is continuing to investigate approaches to address this issue. EPA is also considering for future analyses means of integrating multiple years of meteorological data into the air quality modeling step; this will help address potential uncertainties associated with using a single year’s meteorological data to model conditions across multiple target years.

The modeling of non-point/area sources may also play a role. When compared against the 1990 base year AERMOD run, the average benzene concentration attributed to non-point/area sources in the 2000, 2010, and 2020 *Without-CAAA* runs appears to decrease, despite greater non-point/area source emissions in each of those years (see Appendix B). These results appear to reflect the sensitivity of the air quality modeling to differences between the surrogate data used in the 1990 model run to allocate non-point/area source emissions and the surrogate data from 2000 used in all the future year model runs. This is a potentially significant source of uncertainty; if the 2000 allocation surrogate data *Without-CAAA* allocate area source emissions in such a way that the dispersion model systematically underestimate concentrations from area sources in the *Without-CAAA*

---

[76] This would occur because the benzene concentrations in the *Without-CAAA* scenario are typically higher than those in the *With-CAAA* scenario. If the downward bias is proportional to the concentration, the *Without-CAAA* value would be more significantly underestimated than the *With-CAAA* value, resulting in a smaller than expected difference between the two values.

scenario, our benefits estimates could be underestimated. Because the 2000 allocation is based on more recent data, we believe it is likely more accurate than the 1990 allocation. However, we note that the 2000 allocation surrogate has not yet been validated. Use of a consistent set of validated surrogates is recommended for future assessments.

On the exposure assessment side, there are a limited number of microenvironments included in the HAPEM6 model; as a result, we were unable to estimate benefits expected to occur in certain high-exposure microenvironments such as service stations and homes with attached garages. As a result our benefits may underestimate benefits that occur in these microenvironments. In a supplemental back-of-the-envelope calculation of the magnitude of benefits to those living in homes with attached garages, we estimated benefits of similar magnitude to our primary estimate in 2020. Future analyses would benefit from collecting improved data on the benzene exposures due to attached garages, and from exploring the proportion of benzene exposure risk attributable to indoor sources to provide an overall public health context.

4.2.2 HEALTH BENEFITS MODELING AND VALUATION

Uncertainties related to health benefits modeling and valuation include the following:

- We only quantified health benefits due to avoided cases of leukemia. Other health endpoints associated with benzene exposure that are biologically plausible but lacked sufficient data to quantify a dose-response relationship include other cancers, such as Hodgkin's Lymphoma, and non-Hodgkin's Lymphoma, multiple myeloma, and myeodysplastic syndrome as well as potential non-cancer effects. Therefore, our results do not provide a comprehensive estimate of health benefits from benzene reductions in the Houston area. The SAB (USEPA, 2008b) specifically recommended examining recent studies linking benzene and non-Hodgkin's lymphoma for future benzene benefits analyses.

- We obtained the widest range of benefits from our model (between 0.8 and 7 avoided cases of leukemia) by applying the model to the bounds of the 95 percent confidence interval around our primary dose-response coefficient. Our model is also sensitive to alternative assumptions about dose-response and cessation lag models for benzene-induced leukemia. Sensitivity analyses show that our results can vary by plus 66 percent to minus 81 percent, depending on the choice of cohort study (Pliofilm vs. Chinese Worker), exposure matrix (Crump and Allen versus Paustenbach), health endpoint (total leukemias vs. Acute Myelogenous Leukemia (AML)), or risk/lag model.

- The leukemia cohort studies are based on occupational exposure levels. Extrapolation of the dose-response function to ambient environmental levels requires an assumption of the shape of the function in the observable range. While we have assumed a linear function, as described in Chapter 2 and Appendix C, there is some evidence to suggest the function may be supra-linear; if so, we will have underestimated the benefits of CAAA benzene reductions. Additional research, both epidemiological and toxicological, can help further our understanding of the mode of action of benzene and will help analysts better ascribe probabilities to the alternative functional forms.

- We have applied the relative risk model derived from Crump 1994 to all age groups; however the risk estimates were derived from an occupationally exposed cohort of adults. We may under- or overestimate risk to age groups not included in the cohort if their true relative risk is higher or lower, respectively, than that of the age groups in the worker cohort.[77]
- Application of risk estimates derived from an occupational epidemiological study to the general population typically underestimates risks to that population because the population studied was on average healthier than the general population (i.e., the "healthy worker" effect; Hennekens and Buring, 1987). Because we apply the leukemia risk estimate without adjustment for this effect, the healthy worker effect will tend to bias our results downwards.
- We assumed our linear dose-response model exhibited no threshold (i.e., no exposure level below which no effect would be observed). As discussed in Appendix C, there exists some limited evidence suggesting that a threshold may exist; if the true model exhibits a threshold, our results would be biased upward. The degree of bias would depend on the location of that threshold.
- Our approach for quantifying non-cancer health effects resulting from benzene exposure relied on the RfC reported in IRIS. More recent studies have reported decreased lymphocyte count at benzene concentrations lower than the RfC. Therefore, it is possible that CAAA controls may have resulted in reductions in non-cancer effects in the study population that are not quantified in our analysis.
- Our primary monetized benefit results are highly sensitive to the discount rate applied, because the cessation lag effect delays the full realization of health benefits.[78]
- The VSL value we applied ($7.4 million in 2006$) is a central estimate from a distribution of values obtained from the benefits valuation literature. Use of alternative values from this distribution would scale our monetized benefits accordingly. However, this VSL distribution does not reflect any additional willingness-to-pay to avoid the additional pain and suffering associated with a cancer-related death, and is not included in the pre-mortality morbidity estimate we add to the VSL. To the extent individuals would pay more to avoid cancer-related pain and suffering prior to death, we are underestimating the value of the avoided leukemia cases (i.e., our results do not incorporate a "cancer premium").

---

[77] We used the same relative risk estimates for all groups in this analysis. Because benzene's MOA has not been established at this time, we did not apply the age dependent adjustment factors (ADAFs) recommended in the Supplemental Guidance for Assessing Susceptibility Early-Life Exposures to Carcinogens (USEPA, 2005b) for chemicals with a mutagenic MOA. Early-life adjustments could be explored in a future case study.

[78] Alternative risk models with shorter lags are less sensitive to choice of discount rate, because benefits of exposure reductions will be realized sooner.

Additional studies addressing this issue would significantly benefit HAP analyses, since many of these compounds exhibit carcinogenic effects.

- As noted in Chapter 2, valuation estimates for non-fatal cancers are quite limited. While the approach we employed does build on precedent from past regulatory analyses to generate a willingness-to-pay (WTP) estimate, this estimate consists of two data points, only one of which represents WTP to avoid a case of cancer and neither or which specifically addresses leukemia. Additional research is needed to develop WTP estimates for leukemias and other non-fatal cancers.

4.3 IMPLICATIONS FOR FUTURE ANALYSIS

This case study has demonstrated a benefits methodology that can be used to assess the health impacts of changes in benzene concentrations in an urban area. As EPA moves forward in its development of benefit analysis tools for HAPs, it should consider the potential role of this methodology in more broadly documenting the effects of HAP regulation on health.

In 2001, Agency staff and members of the EPA SAB held a joint workshop to explore the issue of how to best estimate and value the benefits of HAP reductions. The workshop, which included experts in economics, health science, and risk assessment, engendered extensive discussion, but yielded no consensus as to the best methodology. Participants were divided over the use of traditional damage-function approach, such as the one applied in this case study. The SAB workshop report cites a number of obstacles to this approach, including limited, often contradictory, health data; difficulty assessing the effects of multiple exposures; uncertainties in extrapolating from animals to humans and from high doses to low doses; and limited resources to evaluate a large number of chemicals (USEPA SAB, 2002b). The workshop concluded with recommendations for two research directions: one pursuing the demonstration of the damage-function approach for a well-studied HAP and the other pursuing alternative approaches suggested at the meeting, such as assessing the value of HAP regulation as an insurance policy or assessing the value of shifts in the curve of a population's onset of disease (USEPA SAB, 2002b).

This study provides insights into the strengths and limitations of a damage-function approach. Specific strengths of the methodology applied in this case study include:

- It provides a comprehensive assessment of the impact of benzene controls from multiple CAAA Titles on cancer incidence in an urban population;
- It uses a combination of national and local data to develop emissions inventories cost-effectively, which include improved resolution link-level mobile source emissions estimates;
- It assesses exposure using EPA's HAPEM model, which combines air quality modeling output from AERMOD with local activity pattern (e.g., commuting) data to generate both more realistic, age-specific estimates of exposures at the census tract level and probabilistic distributions of exposure that reflect interpersonal variability in exposure;

- It generates health benefit estimates based on central, rather than upper-bound, estimates of cancer potency, which is more appropriate for regulatory analysis;
- It applies a life-table model which allows for the assessment of the CAAA benzene controls on the population over time, using the age-specific HAPEM exposure estimates and local, age-specific baseline incidence rates to generate estimates of local health impacts by census tract;
- It simplifies the consideration of cessation lag by integrating it directly into the life-table model, which uses a damage-function based on weighted exposures; and
- It generates monetized estimates of avoided cancer cases, both fatal and non-fatal, using current EPA guidance on VSL estimates for cancer.
- It uses a modular approach to the analysis, which provides opportunities for scaling the level of complexity of the analysis in accordance with needs and resources.

Specific limitations of the methodology and drawbacks to wider application include:

- The damage-function approach requires both significant resources and extensive data sets to perform local-scale modeling;
- The number of HAPs with a sufficient toxicological database in terms of number and quality of studies and weight of evidence to support this type of health benefits modeling remains limited;
- Use of the model with HAPs other than benzene may require additional effort to estimate a central-estimate dose-response function from available data, as many published toxicological values for other HAPs represent upper bound estimates of potency or reference values that do not allow for quantitative risk assessment;
- The model has not yet been demonstrated for a non-cancer dose-response analysis.
- The critical effects associated with published non-cancer toxicological benchmarks for many HAPs may be difficult to value economically, because while they may serve as an indicator of an adverse biological process, the effects themselves may not necessarily be clinically significant (e.g. increased kidney weights); and
- Air toxics monitoring is more limited than criteria pollutant modeling, making it more difficult to conduct quality control model-to-monitor comparisons in some locations or for certain HAPs.

The drawbacks of applying this model more broadly are essentially the same as those cited in the 2001 workshop, though there have been some positive developments for HAP benefits assessment. For example, EPA's 2005 *Guidelines for Carcinogen Risk Assessment* encourages improved reporting of uncertainty in risk estimates, including central as well as high-end estimates. In addition, since 2002, EPA's IRIS database has updated 23 toxicological summaries, 11 of which were for HAPs.[79] Unfortunately, insufficient data exist for most of these HAPs to assess their carcinogenic potency. One of the updated HAPs - 1,3-butadiene - is classified as carcinogenic to humans and does appear to have a sufficient database to support benefits analysis, including epidemiological results showing a dose-response relationship for leukemias in polymer workers in the U.S. (USEPA, 2007e). 1,3-Butadiene is one of the 12 regional cancer risk drivers identified in EPA's 1999 National Air Toxics Analysis (NATA) analysis (USEPA, 2001c), and therefore may be a good candidate for further analysis using this model.[80]

In order to apply the methodology to a non-carcinogen, additional effort would be required to develop a dose-response function for use with the health effects model. While the resulting function and estimated benefits would be uncertain, there is also significant uncertainty in the true impacts of exposures in a population simply characterized as being above the RfC. Experts have argued for a more parallel treatment of carcinogens and non-carcinogens (e.g., Clewell and Crump, 2005), and a recent paper by Woodruff et al. (2007) illustrated an approach to developing a dose-response model for acrolein, the one HAP identified as a risk driver of non-cancer effects at the national level in EPA’s 1999 NATA.[81]

We believe future case studies should continue to provide both central estimates of population risk (i.e., estimates of cases of adverse health effects avoided) and estimates of individual risk reductions for highly exposed populations. The latter are particularly important, because the impacts of HAP emissions (and emission reductions) can be fairly localized, as seen in the substantial risk reductions in high exposure tracts in Brazoria and Galveston counties.

In an effort to ascertain how our benefits may compare to those estimated from a larger-scale analysis such as EPA's National Air Toxics Assessment (NATA), IEc attempted to conduct a reduced-form benefits analysis of CAAA-related benzene reductions in the three-county study area using benzene concentrations from the 1999 NATA and preliminary draft concentrations from the forthcoming 2002 NATA. However, we found the NATA results to be incompatible with our benefits model, because for many census tracts the NATA results (from both 1999 and 2002) exceeded both the with- and without-CAAA estimates from our case study. While we were unable to conduct a thorough investigation of the causes of these discrepancies, our initial efforts suggest that differences in year 2000 onroad benzene emissions are a contributing factor; additional contributors may be differences in air quality modeling (AERMOD vs. ASPEN), the

---

[79] See http://www.epa.gov/iris/whatsnew.htm and http://www.epa.gov/iris/whatsnewarch.htm for updated profiles. The 11 HAPs were vinylidene chloride (1,1-dichloroethylene); phenol; 1,3-butadiene; xylenes, benzene, methylisobutylketone, acrolein, toluene, hexane, phosgene, and 2,2,4-trimethylpentane.

[80] The NATA study identifies regional risk driver as carcinogens to which at least one million people are exposed at a risk level greater than 10 in one million or at least 10,000 people are exposed at a risk level greater than 100 in one million. See http://www.epa.gov/ttn/atw/nata1999/ for the full list of cancer and non-cancer risk drivers.

[81] A national risk driver for non-cancer effects, as defined in the 1999 NATA, is a HAP for which at least 25 million people are exposed at levels above EPA's reference concentration. The study also identified 16 HAPs as regional drivers of non-cancer risk, defined as HAPs for which at least 10,000 people are exposed above EPA's reference concentration.

apparent gaps in our 1990 base year benzene emission inventory that contribute to underestimated benzene concentrations in 2000 and likely throughout the study (discussed earlier), and uncertainty in the benzene emissions growth factors used to generate the with- and without-CAAA scenarios for this case study. Future HAP analyses, whether at the urban-scale or using NATA would benefit from further investigations of the differences in these approaches, the associated uncertainties and data gaps and their potential impact on results.

Due to difficulties in applying the case study approach on a national scale or to extending it to other air toxics, which may have a limited epidemiological database, the SAB in its review of this case study (USEPA, 2008b) suggested that EPA also consider integrated multi-pollutant approaches to estimate the benefits of air toxics regulations. For example, OAR’s Risk and Technology Review (RTR) program, which evaluates air toxics risk by source category. Another option could be the emerging 3D air quality modeling work that can include individual air toxics so that HAPs do not need to be modeled separately.

In conclusion, the methodology presented in this case study can serve as a useful tool in EPA's evolving HAP benefit assessment strategy. Determining where this approach best fits within that strategy will require additional analysis and evaluation to determine the added value of the detailed, urban-scale approach, as well as potential pool of HAPs suitable for assessment via the damage-function approach for cancer and/or non-cancer effects.

REFERENCES

Adegoke, O.J., Blair, A., et al. (2003). Occupational history and exposure and the risk of adult leukemia in Shanghai. *Annals of Epidemiology* 13:485-494.

Allen, D. and Durrenberger, C. (2003). Accelerated Science Evaluation of Ozone Formation in the Houston-Galveston Area: Emission Inventories, Version 3. Texas Commission on Environmental Quality (formerly Texas Natural Resource Conservation), Technical Analysis Division. February, 2003. Available at http://www.utexas.edu/research/ceer/texaqsarchive/accelerated.htm

Batelle, (2003). *Estimate background concentrations for the National-Scale Air Toxics Assessment*. Technical Report. Prepared for U.S. Environmental Protection Agency. Contract No. 68-D-02-061.

Bloemen, L.J., Youk, A., et al. (2004). Lymphohaematopoietic cancer risk among chemical workers exposed to benzene. *Occupational and Environmental Medicine* 61:270-274.

California Environmental Protection Agency (2001). *Public Health Goal for benzene in drinking water*. Office of Environmental Health Hazard Assessment.

Centers for Disease Control (CDC) WONDER online database. http://wonder.cdc.gov/.

Clewell, H.J. and Crump, K.S. (2005). Quantitative estimates of risk for noncancer endpoints. *Risk Analysis* 25(2):285-9.

Cohen, J.; Cook, R.; Bailey, C.R.; Carr, E. (2005). Relationship between motor vehicle emissions of hazardous pollutants, roadway proximity, and ambient concentrations in Portland, Oregon. *Environ Modelling & Software* 20: 7-12.

Collins, J.J., Ireland, B.K., et al. (2003). Lymphohaematopoeitic cancer mortality among workers with benzene exposure. *Occupational and Environmental Medicine* 60:676-679.

Cook, Richard. EPA/Office of Transportation and Air Quality. *Compare9902.xls* [electronic file]. Transmitted to Henry Roman of Industrial Economics, Inc. via Jeneva Craig of EPA/OPAR, April 18, 2007.

Costantini, A.S., Quinn, M., et al. (2003). Exposure to benzene and risk of leukemia among shoe factory workers. *Scandanavian Journal of Work and Environmental Health* 29(1):51-59.

Crump, K.S. (1996). Risk of benzene-induced leukemia predicted from the Pliofilm Cohort. *Environmental Health Perspectives* 104 (Suppl 6): 1437-1441.

Crump, K.S. (1994). Risk of benzene-induced leukemia: A sensitivity analysis of the Pliofilm Cohort with additional follow-up and new exposure estimates. *Journal of Toxicology and Environmental Health* 42:219-242.

Crump, K.S., and Allen, B.C. (1984). *Quantitative estimates of risk of leukemia from occupational exposure to benzene*. US Department of Labor, Washington DC (OSHA Docket H-059b, Exhibit 152, Annex B).

Dosemeci, M., Li, G.-L., et al. (1994). Cohort study among workers exposed to benzene in China: II. Exposure assessment. *American Journal of Industrial Medicine* 26:401-411.

Energy Information Administration (2005a). *State Energy Consumption, Price, and Expenditure Estimates (SEDS)*. U.S. Department of Energy, Washington, DC, Texas data retrieved on July 20, 2005 from http://www.eia.doe.gov/emeu/states/state.html?q_state_a=tx&q_state=TEXAS.

Energy Information Administration (2005b). *Distribution and Production of Oil and Gas Wells by State.* U.S. Department of Energy, Washington, DC, Texas data retrieved on July 22, 2005 from http://www.eia.doe.gov/pub/oil_gas/petrosystem/petrosysog.html.

Energy Information Administration (2005c). *Sales of Distillate Fuel Oil by End Use.* U.S. Department of Energy, Washington, DC, Texas data retrieved on July 20, 2005 from http://tonto.eia.doe.gov/dnav/pet/pet_cons_821dst_dcu_STX_a.htm.

Energy Information Administration (2005d). *Annual Energy Outlook 2005 With Projections to 2025*, DOE/EIA-0383(2005), U.S. Department of Energy, Washington, DC, February 2005.

E.H. Pechan and Associates (2006). *Section 812 Clean Air Act Cost-Benefit Study: Air Toxics Case Study Benzene Emission Reductions In Houston*. Prepared for Industrial Economics, Inc., February.

European Commission Directorate-General XII Science, Research and Development. Common Annexes of the ExternE National Implementation Reports (1998), downloaded 6/1/07 from: http://www.externe.jrc.es/reports.html.

Fischer, P.H.; Joek, G.; van Reeuwijk, H.; et al. (2000). Traffic-related differences in outdoor and indoor concentrations of particle and volatile organic compounds in Amsterdam. *Atmos Environ* 34: 3713-3722.

Glass, D.C., Gray, C.N., et al. (2003). Leukemia risk associated with low-level benzene exposure. *Epidemiology* 14(5):569-577.

Glen, G., Lakkadi, Y., et al. (1997). *Development of NERL/CHAD: The National Exposure Research Laboratory Consolidated Human Activity Database.* EPA Contract No. 68-D5-0049.

Gordon, S.M.; Callahan, P.J.; Nishioka, M.G.; Brinkman, M.C.; O'Rourke, M.K.; Lebowitz, M.D.; Moschandreas, D.J. (1999). Residential environmental measurements in the National Human Exposure Assessment Survey (NHEXAS) pilot study in Arizona: preliminary results for pesticides and VOCs. *J. Expos. Analysis. Environ. Epidemiol.* 9: 456-470.

Graham, L.A., Noseworthy, L. (2004). Contribution of vehicle emissions from an attached garage to residential indoor air pollution levels. *Journal of the Air and Waste Management Association* 54: 563-584.

Guenel, P., Imbernon, E., et al. (2002). Leukemia in relation to occupational exposures to benzene and other agents: A case-control study nested in a cohort of gas and electric utility workers. *American Journal of Industrial Medicine* 42:87-97.

Hammitt JK, Liu J-T (2004). Effects of disease type and latency on the value of mortality risk. *Journal of Risk and Uncertainty* 28:73-95.

Hao, J.; He, D.; Wu, Y.; Fu, L.; He, K. (2002). A study of the emission and concentration distribution of vehicular pollutants in the urban area of Beijing. *Atmospheric Environment* 34: 453-465.

Hayes, R.B., Yin, S.-N., et al. (1997). Benzene and the dose-related incidence of hematologic neoplasms in China. *Journal of the National Cancer Institute* 89(14):1065-1071.

Hennekens, C.H. and Buring, J.E. (1987) *Epidemiology in Medicine*. Little, Brown, and Company: Boston.

ICF International. (2007). The *HAPEM6 User's Guide: Hazardous Air Pollutant Exposure Model, Version 6*. Prepared for the Office of Air Quality Planning and Standards, US EPA, Research Triangle Park, NC.

Ilgen, E.; Karfich, N.; Levsen, K.; et al. (2001). Aromatic hydrocarbons in the atmospheric environment: part I. Indoor versus outdoor sources, the influence of traffic. *Atmos Environ* 35: 1235-1252.

Industrial Economics, Inc. (2003). *Benefits and Costs of the Clean Air Act 1990-2020: Revised Analytical Plan for EPA's Second Prospective Analysis*. Prepared for EPA's Office of Policy Analysis and Review, Washington, DC.

Industrial Economics, Inc. (2005). Memorandum to Jenny Craig, EPA Office of Policy Analysis and Review. "Benzene Health Effects Literature Review." July 5.

Ireland, B., Collins, J.J., et al. (1997). Cancer mortality among workers with benzene exposure. *Epidemiology* 8:318-320.

Jo, W.; Kim, K.; Park, K.; et al. (2003). Comparison of outdoor and indoor mobile source-related volatile organic compounds between low- and high-floor apartments. *Environ Res* 92: 166-171.

Kleinman, L.I., Daum, P.H., Imre, D., Lee, Y.N., Nunnermacker, L.J., Springston, S.R., Weinstein-Lloyd, J., Rudolph, J., 2002. Correction to "Ozone production rate and hydrocarbon reactivity in 5 urban areas: A cause of high ozone concentration in Houston." Geophys. Res. Lett. 30 (12), 1639.

Kochi, I., Hubbell, B., et al. (2006). An Empirical Bayes Approach to Combining Estimates of the Value of Statistical Life for Environmental Policy Analysis. *Environmental and Resource Economics* 34: 385-406.

Kwon, J. (2005). *Development of a RIOPA database and evaluation of the effect of proximity on the potential residential exposure to VOCs from ambient sources. Rutgers, the State University of New Jersey and University of Medicine and Dentistry of New Jersey*. PhD dissertation. This document is available in Docket EPA-HQ-OAR-2005-0036.

Lan, Q, Zhang, L., et al. (2004). Hematotoxicity in workers exposed to low levels of benzene. *Science* 306: 1774-1776.

MedlinePlus Website (2008). A service of the U.S. National Library of Medicine and the National Institutes of Health. http://medlineplus.gov/

National Cancer Institute (NCI). (2008). Leukemia Booklet (NIH Publication No. 08-3775). http://www.cancer.gov/cancertopics/wyntk/leukemia.

National Research Council (NRC). (1988). *Health Risks of Radon and Other Internally Deposited Alpha-Emitters (BEIR IV).* Washington, DC: National Academy Press.

Magat, W.A., Viscusi, W.K., et al. (1996). A reference lottery metric for valuing health. *Management Science* 42:1118-1130.

Paustenbach, D.J., Price, P.S., et al. (1992). Reevaluation of benzene exposure for the Pliofilm (rubberworker) Cohort (1936-1976). *Journal of Toxicology and Environmental Health* 36:177-231.

Payne-Sturges, D.C., Burket, T.A., et al. (2004). Personal exposure meets risk assessment: A comparison of measured and modeled exposures and risks in an urban community. *Environmental Health Perspectives* 112: 589-598.

Qu, Q., Shore, R., et al. (2002). Hematological changes among Chinese workers with a broad range of benzene exposures. *American Journal of Industrial Medicine* 42:275-285.

Rinsky, R.A., Hornung, R.W., et al. (2002). Benzene exposure and hematopoietic mortality: A long-term epidemiologic risk assessment. *American Journal of Industrial Medicine* 42:474-480.

Rinsky, R.A., Young, R.J., et al. (1981). Leukemia in benzene workers. *American Journal of Industrial Medicine* 2(3):217-45.

Rinsky, R.A., Smith, A.B., et al. (1987). Benzene and leukemia: An epidemiologic risk assessment. *New England Journal of Medicine* 316:1044-1050.

Rothman, N., Li, G.-L., et al. (1996a). Hematotoxicity among Chinese workers heavily exposed to benzene. *American Journal of Industrial Medicine* 29:236-246.

Rothman, N, Smith, M.T., et al. (1996b). An epidemiologic study of early biologic effects of benzene in Chinese workers. *Environmental Health Perspectives* 104(Suppl 6):1365-1370.

Rothman, N., Smith, M.T., et al. (1997). Benzene poisoning, a risk factor for hematological malignancy, is associated with the NQ01609C-T mutation and rapid fractional excretion of chlorzoxazone. *Cancer Research* 57:2839-2842.

Rusthton, L. and Romaniuk, H. (1997). A case-control study to investigate the risk of leukaemia associated with exposure to benzene in petroleum marketing and distribution workers in the United Kingdom. *Occupational and Environmental Medicine* 54(3):152-166.

Ryerson, T.B. Trainer, M., Angevine, W.M., Brock, C.A., Dissly, R.W., Fehsenfeld F.C., Frost, G.J., Goldan, P.D., Holloway, J.S., Hubler, G., Jakoubek, R.O., Kuster, W.C., Neumann, J.A., Nicks, D.K., Parrish, D.D., Roberts, J.M., Sueper, D.T., (2003). Effect of petrochemical industrial emissions of reactive alkenes and NOx on tropospheric ozone formation in Houston, Texas. *Journal of Geophysical Research—Atmospheres 108 (D8), Art. No. 4249.*

Sapkota, A.; Buckley, T.J. (2003). The mobile source effect on curbside 1,3-butadiene, benzene, and particle-bound polycyclic aromatic hydrocarbons assessed at a tollbooth. *J Air Waste Manage Assoc* 53: 740-748.

Schlapia, A.; Morris, S.S. (1998). Architectural, behavioral and environmental factors associated with VOCs in Anchorage homes. *Proceedings of the Air & Waste Management Association's 91st Annual Meeting & Exhibition*. Paper no. 98-A504.

Schnatter, R.A., Armstrong, T.W., et al. (1996). Lymphohaematopoietic malignancies and quantitative estimates of exposure to benzene in Canadian petroleum distribution workers. *Occupational and Environmental Medicine* 53(11):773-781.

Science Policy Council (SPC) (2000). *Handbook for Non-Cancer Health Effects Valuation*. U.S. Environmental Protection Agency, Washington, DC, December.

Silver, S.R., Rinsky, R.A., et al. (2002). Effect of follow-up time on risk estimates: A longitudinal examination of the relative risks of leukemia and multiple myeloma in a rubber hydrochloride cohort. *American Journal of Industrial Medicine* 42:481-489.

Skov, H.; Hansen, A.B.; Lorenzen, G.; et al. (2001). Benzene exposure and the effect of traffic pollution in Copenhagen, Denmark. *Atmos Environ* 35: 2463-2471.

Stein, A.F., Isakov, V., Godowitch, J., and Draxler, R.R. (2007). A hybrid modeling approach to resolve pollutant concentrations in an urban area. *Atmos Environ* 41 (40): 9410-9426.

Stein, B., Walker, D., Cook, R. and Bailey, C. (2002). *Link-Based Calculation of Motor Vehicle Air Toxin Emissions Using MOBILE6.2*. Metro Planning Department, Portland, OR, and U.S. Environmental Protection Agency, Office of Transportation and Air Quality, Ann Arbor, MI.

Sorahan, T., Kinlen, L.J., et al. (2005). Cancer risks in a historical UK cohort of benzene exposed workers. *Occupational and Environmental Medicine* 62:231-236.

Thurman, James, CSC, Inc. *onroad_comparisons.xls; stationary_comparisons.xls; nonroad_comparisons.xls* [electronic files]. Transmitted to Henry Roman of Industrial Economics, Inc. via Neal Fann of EPA/OAQPS, May 10, 2006.

Tsuge, T., A. Kishimoto and K. Takeuchi, 2005. A Choice Experiment Approach to the Valuation of Mortality. *Journal of Risk and Uncertainty*, 31(1): 73-95, July.

Turtletaub, K.W., Mani, C. (2003). *Benzene metabolism in rodents at doses relevant to human exposure from urban air*. Health Effects Institute Research Report 113, Boston, MA.

U.S. Environmental Protection Agency. *1999 National Emission Inventory Documentation and Data - Final Version 3.0*. From: http://www.epa.gov/ttn/chief/net/1999inventory.html.

U.S. Environmental Protection Agency. *2002 National Emissions Inventory Data & Documentation, Version 2*. http://www.epa.gov/ttn/chief/net/2002inventory.html.

U.S. Environmental Protection Agency (1996). *Proposed Guidelines for Carcinogen Risk Assessment*. Office of Research and Development, Washington, DC. EPA/600/P-92/003C.

U.S. Environmental Protection Agency (1998). *Carcinogenic Effects of Benzene: An Update.* Office of Research and Development, Washington, DC. EPA/600/P-97/001F.

U.S. Environmental Protection Agency, (1999a) *The Benefits and Costs of the Clean Air Act: 1990 to 2010*. EPA-410-R-99-001.

U.S. Environmental Protection Agency, (1999b). *Cost of Illness Handbook.* http://www.epa.gov/oppt/coi/pubs/I_1.pdf.

U.S. Environmental Protection Agency (2000). *An SAB Report on EPA's White Paper Valuing the Benefits of Fatal Cancer Risk Reductions*. Science Advisory Board. EPA-SAB-EEAC-00-013.

U.S. Environmental Protection Agency (2001a). *Arsenic Rule Benefits Analysis: An SAB Review*. Prepared by the Arsenic Rule Benefits Review Panel (ARBRP) of the US EPA Science Advisory Board (SAB). Washington, DC. EPA-SAB-EC-01-008.

U.S. Environmental Protection Agency. (2001b). *Review Of The Draft Analytical Plan For EPA's Second Prospective Analysis - Benefits And Costs Of The Clean Air Act 1990-2020.* Science Advisory Board, Washington, D.C. EPA-SAB-COUNCIL-ADV-01-004.

U.S. Environmental Protection Agency, (2001c). *National-Scale Air Toxics Assessment for 1999*. http://www.epa.gov/ttn/atw/nata1999/.

U. S. Environmental Protection Agency, (2002a). *Example Application of Modeling Toxic Air Pollutants in Urban Areas*. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/R-02-003. http://www.epa.gov/scram001/guidance/guide/uatexample.pdf

U. S. Environmental Protection Agency (2002b) *Workshop on the Benefits of Reductions in Exposure to Hazardous Air Pollutants: Developing Best Estimates of Dose-Response Functions, An SAB Workshop Report of an EPA/SAB Workshop.* Science Advisory Board, Washington, D.C. EPA-SAB-EC-WKSHP-02-001.

U.S. Environmental Protection Agency (2004a) Office of Transportation and Air Quality, NONROAD2004 Model, [Computer software], Ann Arbor, MI, May 2004. Available at: http://www.epa.gov/otaq/nonrdmdl.htm.

U.S. Environmental Protection Agency (2004b). *User's Guide for the AMS/EPA Regulatory Model – AERMOD*. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-001. http://www.epa.gov/scram001/7thconf/aermod/aermodugb.pdf

U.S. Environmental Protection Agency (2004c). *User's Guide for the AERMOD Meteorological Preprocessor (AERMET)*. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-002. http://www.epa.gov/scram001/7thconf/aermod/aermetugb.pdf

U. S. Environmental Protection Agency (2004d). *User's Guide for the Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP, Version 3.0)*. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-006. http://www.epa.gov/scram001/dispersion_related.htm - ems-hap

U.S. Environmental Protection Agency. (2004e). *Advisory on Plans for Health Effects Analysis in the Analytical Plan for EPA's Second Prospective Analysis - Benefits and Costs of the Clean Air Act, 1990-2020, Advisory by the Health Effects Subcommittee of the Advisory Council for Clean Air Compliance Analysis*. Science Advisory Board, Washington, D.C. EPA-SAB-COUNCIL-ADV-04-002.

U.S. Environmental Protection Agency (2005a). Office of Transportation and Air Quality, NMIM Software (NMIM20050311) and NMIM County Database (NCD20050427), Ann Arbor, MI, April 2005, available at: http://www.epa.gov/otaq/nmim.htm

U.S. Environmental Protection Agency (2005b). *Guidelines for carcinogen risk assessment*. Risk Assessment Forum, Washington, D.C. EPA/630/P-03/001F.

U.S. Environmental Protection Agency (2005c). *2002 EPA National Emission Inventory Point Source Database (Draft)*. Office of Air Quality Planning and Standards, Emission Inventory Group, http://www.epa.gov/ttn/chief/net/2002inventory.html#point, February 2005.

U.S. Environmental Protection Agency (2005d). *Documentation for the 1990 National Emission Inventory for Hazardous Air Pollutants*. Emission Inventory Group, Research Triangle Park, NC, November 15, 2005.

U.S. Environmental Protection Agency (2005e). Technology Transfer Network, http://www.epa.gov/ttn/chief/emch/projection/emshap30.html, Clearinghouse for Emission Inventories and Factors, Research Triangle Park, NC, accessed July 2005.

U.S. Environmental Protection Agency (2006). *Regulatory Impact Analysis for the 2006 National Ambient Air Quality Standards for Particle Pollution.* http://www.epa.gov/ttn/ecas/ria.html

U.S. Environmental Protection Agency (2007a). *Air Quality and Exposure Modeling in Support of Section 812 Benzene Case Study: Methodology and Results .Draft.*

U.S. Environmental Protection Agency (2007b). Integrated Risk Information System Toxicological profile for benzene. Searched June, 2007.

U.S. Environmental Protection Agency (2007c). Integrated Risk Information System Glossary of Terms. Searched June, 2007.

U.S. EPA. (2007d). *Control of Hazardous Air Pollutants from Mobile Sources: Regulatory Impact Analysis*. Office of Transportation and Air Quality. EPA420-R-07-002.

U.S. Environmental Protection Agency (2007e). Integrated Risk Information System Toxicological profile for 1,3-butadiene. Searched August, 2007.

U.S. Environmental Protection Agency (2008a). *BenMAP User's Manual*. Prepared by Abt Associates, Inc. for the Office of Air Quality Planning and Standards, September.

U.S. Environmental Protection Agency (2008b). Letter from James K. Hammitt, Chair, Advisory Council on Clean Air Compliance Analysis to Stephen L. Johnson, Administrator of the U.S. EPA re: Benefits of Reducing Benzene Emissions in Houston, 1990-2020, July 11.

Van Houtven, G., Sullivan, M.B., and Dockins, C. (2008). Cancer premiums and latency effects: A risk tradeoff approach for valuing reductions in fatal cancer risks. *Journal of Risk and Uncertainty* 36: 179-199.

Viscusi, W. Kip, Magat, Wesley, A. and Huber, Joel. (1991). Pricing environmental health risks: survey assessments of risk-risk and risk-dollar tradeoffs for chronic bronchitis. *Journal of Environmental Economics and Management* 21:32-51.

Woodruff T.J., Wells E.M., Holt E.W., Burgin D.E., Axelrad D.A. (2007). Estimating risk from ambient concentrations of acrolein across the United States. *Environ Health Perspect.* 115(3):410-5.

Woods and Poole Economics, Inc. (2001). Population by Single Year of Age CD. CD-ROM. Woods & Poole Economics, Inc. Washington, D.C.

Wong, O. (1995). Risk of acute myeloid leukaemia and multiple myeloma in workers exposed to benzene. *Occupational and Environmental Medicine* 52:380-384.

Yin, S.-N., Hayes, R.B., et al. (1996). An expanded cohort study of cancer among benzene-exposed workers in China. *Environmental Health Perspectives* 104(Suppl 6):1339-1341.

APPENDIX A:

BENZENE EMISSIONS MODELING REPORT

SECTION 812 CLEAN AIR ACT COST-BENEFIT STUDY
AIR TOXICS CASE STUDY
BENZENE EMISSION REDUCTIONS IN HOUSTON

Prepared for:

Industrial Economics, Inc.
2067 Massachusetts Avenue
Cambridge, MA 02140

Prepared by:

E.H. Pechan & Associates, Inc.
5528-B Hempstead Way
Springfield, VA 22151

March 2009

EPA Contract No. EP-D-04-006
Work Assignment No. 2-5 and 4-6
Pechan Report No. 09.03.001/9221.406

CONTENTS

TABLES

Page

[This page intentionally left blank.]

ACRONYMS AND ABBREVIATIONS

AEO 2005	*Annual Energy Outlook 2005*
ATP	anti-tampering program
CAAA	Clean Air Act Amendments of 1990
CAIR	Clean Air Interstate Rule
CAMD	Clean Air Markets Division
CNG	compressed natural gas
DOE	U.S. Department of Energy
EGAS	Economic Growth Analysis System
EGU	electricity generating unit
EIA	Energy Information Administration
EPA	U.S. Environmental Protection Agency
ETS/CEM	emission tracking system/continuous emissions monitoring
HAP	hazardous air pollutant
I/M	inspection and maintenance
LPG	liquefied petroleum gas
LTO	landing and takeoff
MACT	maximum achievable control technology
NAICS	North American Industrial Classification System
NEI	National Emission Inventory
NEMS	National Energy Modeling System
NESHAP	National Emission Standards for Hazardous Air Pollutants
NMIM	National Mobile Inventory Model
NO_x	oxides of nitrogen
NSPS	New Source Performance Standards
NTI	National Toxics Inventory
OTAQ	Office of Transportation and Air Quality
Pechan	E.H. Pechan & Associates, Inc.
PM_{10}	particulate matter of 10 microns or less
$PM_{2.5}$	particulate matter with an aerodynamic diameter of 2.5 microns or less
RVP	Reid vapor pressure
RWC	residential wood combustion
S/L	State and local
SAB	Science Advisory Board
SCC	Source Classification Code
S-I	spark ignition
SIC	Standard Industrial Classification
SIP	State Implementation Plan
SO_2	sulfur dioxide
SOCMI	Synthetic Organic Chemical Manufacturing Industry
TCEQ	Texas Commission on Environmental Quality
tpy	tons per year
TRI	Toxics Release Inventory

TTI	Texas Transportation Institute
VOC	volatile organic compound
WSC	West South Central

CHAPTER I. INTRODUCTION AND SUMMARY

A. BACKGROUND

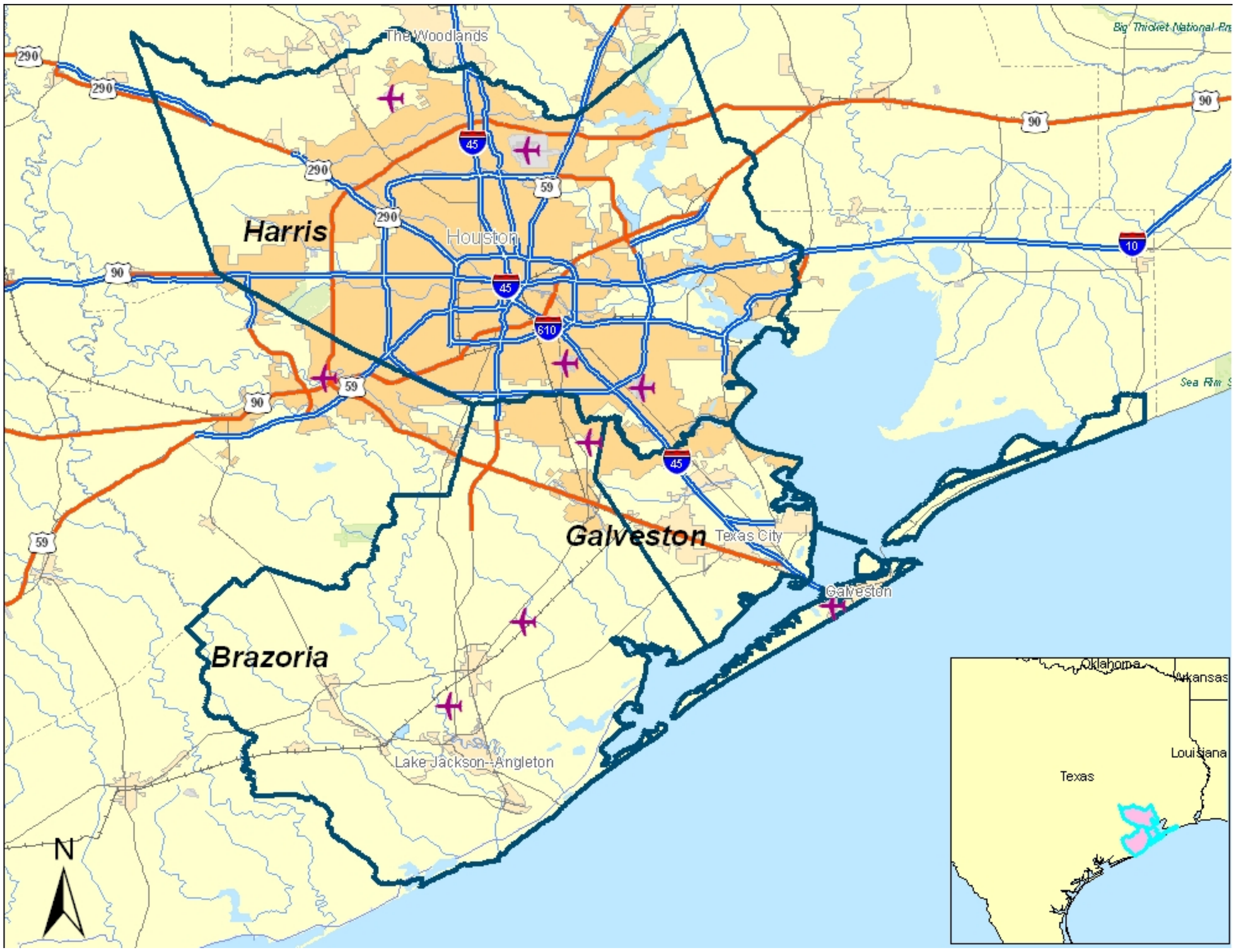
The purpose of the study described in this report is to prepare the benzene emission estimates portion of a hazardous air pollutant (HAP) benefits assessment to accompany the main criteria air pollutant analysis in the second section 812 prospective study. The U.S. Environmental Protection Agency (EPA) Science Advisory Board (SAB) advised EPA to select the representative HAP for which to perform a prototype 812 analysis. The SAB recommended benzene in part because of the wealth of available national ambient concentration data from monitors. The SAB feels that an 812 analysis using the available benzene data would:

- Identify limitations and gaps in the database;
- Provide an estimate of the uncertainties in the analyses and perhaps provide a reasonable lower bound on potential health benefits from control; and
- Provide a scientific basis for deciding whether there is merit in pursuing a greater ability to assess the benefits of air toxics.

EPA's response to these SAB comments was to undertake a metropolitan scale analysis of the benefits of Clean Air Act Amendments of 1990 (CAAA) controls on benzene emissions. This scale allows both a more rigorous analytical effort and the opportunity to build on previous EPA modeling efforts for benzene. The Houston-Galveston area was selected for the case study. The geographic boundary of the study area includes Brazoria, Galveston, and Harris Counties. The original analytical plan for this case study proposed that the case study area be limited to Harris County. The SAB recommended that a larger study area be used. The 1999 EPA National Emission Inventory (NEI) emission estimates for the Houston-Galveston metropolitan area were reviewed and it was found that three counties contained 99 percent of the metropolitan area's point source benzene emissions (Brazoria, Galveston, and Harris Counties). Therefore, it was recommended that the study area include these three counties.

Figure I-1 provides a map of the study area, with the county boundaries outlined for Brazoria, Galveston, and Harris counties.

Figure I-1. Houston-Galveston Benzene Case Study Area



This report is organized by sector. There are four major sectors in this analysis: (1) on-road vehicles, (2) nonroad engines/vehicles, (3) point sources, and (4) nonpoint sources. (Nonpoint sources were formerly called area sources, or area wide sources.) The emission estimation methods and results for each sector are described in a separate chapter. The nonroad engines/vehicles sector typically includes emissions from commercial marine vessels, aircraft, and locomotives. In this report, those sources are included in the nonpoint source analysis, and the non-road sector discussion is limited to the engines and vehicles that are included in EPA's NONROAD model.

As with the main criteria pollutant portion of the section 812 analysis, the analysis years are 1990, 2000, 2010, and 2020. Both *with*- and *without-CAAA* scenarios are evaluated for 2000, 2010, and 2020. The base year is 1990 because that is the year when the CAAA were enacted. Table I-1 summarizes the scenario years that are addressed in this report. Table I-2 summarizes the CAAA provisions that affect benzene emissions in the HGB area during the study period.

Table I-1. Scenarios and Years Addressed in This Study

	Scenario Years						
Sector	**1990**	**2000** ***with-CAAA***	**2000** ***without-CAAA***	**2010** ***with-CAAA***	**2010** ***without-CAAA***	**2020** ***with-CAAA***	**2020** ***without-CAAA***
On-road	✓	✓	✓	✓	✓	✓	✓
Non-road	✓	✓	✓	✓	✓	✓	✓
Point	✓	✓	✓	✓	✓	✓	✓
Nonpoint	✓	✓	✓	✓	✓	✓	✓

For on-road vehicles, the primary CAAA provisions that reduce benzene emissions in the 1990 to 2000 period result from changes in fuel program parameters (via Federal reformulated gasoline [RFG]) and the H-G area enhanced vehicle emissions inspection program. Reductions also result between 1990 and 2000 from vehicle controls in place prior to passage of the 1990 CAAA due to fleet turnover. Tier 1 emission standards reduce VOC emissions (and benzene) somewhat by 2000, but are much less significant than RFG and inspection and maintenance (I/M) programs. By 2010, a greater portion of the emission reductions (29 percent) is attributable to greater phase-in of the Tier 1 VOC emission standards, plus National Low-Emission Vehicle (NLEV) and Tier 2 standards. Vehicle emission standards that affect the light-duty vehicle fleet have a more significant effect on benzene than those affecting heavy-duty diesels.

Off-road engine and vehicle benzene emission changes observed in this analysis are driven by Federal nonroad equipment standards that reduce VOC emissions from spark ignited nonroad engines. This source category is dominated by lawn and garden equipment, so the Federal emission standards for this equipment type dominate the off-road benzene equipment path.

Point source benzene emissions in the HGB area are dominated by the chemical industry, petroleum refining, and fuel storage and distribution. The point source benzene emissions path in the study area is dominated by the Federal MACT standards (Title III) and local VOC control measures (Title I-Nonattainment Provisions) that required the petrochemical facilities in the area to reduce HAP and/or VOC emissions in the years between 1990 and 2000. Where source categories show expected benzene emission decreases between 2002 and 2010 in the with-CAAA scenario, those reductions are attributable to post-2002 MACT standard implementation under Title III. With the strong expected growth in the Gulf Coast area petrochemical industry, absent future VOC or HAP regulation, benzene emissions from point sources are estimated to increase after 2010.

For non-point sources, the most important regulatory driver of future year benzene emissions path is State regulations (Title I-Nonattainment Provisions) affecting sources within the 1-hour ozone nonattainment area. There are also some expected benzene emission reductions in the study period as New Source Performance Standard (NSPS) certified woodstoves replace non-certified stoves.

Table I-2. Benzene Case Study with-CAAA Scenario Summary, by Title

Title I	Any effects of Title I will be expressed through State Implementation Plan (SIP) requirements, such as (enhanced) I/M programs, transportation control measures, and other VOC controls. These requirements are dependent on the ozone nonattainment status of the case study area(s).
Title II	**Tailpipe Standards** **On-road** Tier 1 Standards (phased in 1994 to 1997) NLEV program – voluntary bridge between Tier 1 and Tier 2 Tier 2 Standards take effect in 2004 Heavy Duty Engine/Diesel Fuel Rule – New emission standards – 2007 model year, new fuel standards 2006 **Non-road** Federal Phase I and II compression ignition (CI) engine standards Federal Phase I and II spark ignition (SI) engine standards Federal locomotive standards Federal commercial marine vessel standards Federal recreational marine vessel standards **Evaporative Emissions** State II Vapor Recovery Systems (Section 182) Onboard Refueling Vapor Recovery (Section 202; 1998 model year and on) Evaporative Test Procedure **Fuel Regulations** RFG Standards (1995 on) Phase II – (2000-present) – benzene requirements essentially unchanged Summertime Volatility Requirements for Gasoline (Phase II – 1992 on) Fuel Sulfur Limits *2007 Mobile Source Air Toxics (MSAT) Rule NOT included.
Tier III	**MACT Standards** The With-CAAA scenario included MACT standards that would be expected to have a significant effect on future-year benzene emissions in the Houston area. These standards include: Oil and Natural Gas Production: 7-Year MACT Petroleum Refineries: 4-Year MACT Gasoline Distribution: 4-Year MACT Pulp and Paper Production: 7-Year MACT Municipal Landfills: 10-Year MACT Natural Gas Transmission and Storage: 10pYear MACT Publicly Owned Treatment Works (POTW) Emissions: 7-Year MACT Coke Ovens: Pushing, Quenching, & Battery Stacks: 4-Year MACT Synthetic Organic Chemical Manufacturing Industry Hazardous Organic NESHAP (SOCMI HON): 2-year MACT

B. RESULTS

Table I-3 summarizes the Houston-Galveston study area benzene emission results by sector for the scenarios, sectors, and years evaluated. This table shows that there have been significant CAAA-attributable benzene emission reductions in the Houston-Galveston study area since 1990. A large fraction of these CAAA benefits occurred in the period between 1990 and 2000. For the point and nonpoint source sectors, the benzene emission reductions during this 1990 to 2000 period are largely attributable to Federal maximum achievable control technology (MACT) emission standards, and local volatile organic compound (VOC) measures in the 1-hour ozone attainment plan. The most significant benzene emission reductions in these sectors are from the chemical manufacturing and petroleum refining industries. The emissions path for point and nonpoint benzene sources in the *with-CAAA* scenario is stable or slightly increasing through the 2000 to 2020 period.

Table I-3 also shows that there have been significant CAAA-associated benefits in reducing on-road vehicle emitted benzene in the 1990 to 2000 period, as well. In the Houston-Galveston area, these benefits are primarily attributable to the Federal reformulated gasoline program as well as Federal emission standards that reduced exhaust and evaporative VOCs and benzene emissions. Tier 2 emission standards and associated requirements that lower the sulfur content of gasoline are a significant factor in achieving further benzene emission reductions from onroad vehicles between 2000 and 2020. Chapter V provides more information about this onroad vehicle analysis.

Table I-3. Houston-Galveston Benzene Emissions Summary (tons per year [tpy])

Sector	1990	2000 *without-CAAA*	2000 *with-CAAA*	2010 *without-CAAA*	2010 *with-CAAA*	2020 *without-CAAA*	2020 *with-CAAA*
Point/Nonpoint	5,409	6,532	1,230	6,699	1,258	7,702	1,440
Nonroad	740	900	567	1,127	354	1,351	360
Onroad vehicles	2,375	1,541	762	1,449	328	1,988	282
Total	8,524	8,973	2,559	9,275	1,940	11,041	2,082

Federal nonroad emission standards also produce benzene emission reductions in the Houston-Galveston area. These emission reductions are less than those observed for on-road vehicles in part because the 1990 emissions are much lower, and because many nonroad emission standards were set after the Tier 1 on-road vehicle emission standards took place. Therefore, the biggest differences between *with-* and *without-CAAA* benzene emissions in the study area for nonroad engines/vehicles occur in 2010 and 2020.

The other important observation from Table I-3 is that the relative importance of the four major sectors is changing with time as control programs reduce nonroad and on-road benzene emissions post-2000.

C. UNCERTAINTIES

The NARSTO assessment of the strengths and weaknesses of current emission inventories provides a qualitative ranking of confidence levels in U.S. emissions inventories (NARSTO, 2005). For hazardous air pollutants, NARSTO provides the following confidence levels by source sector:

Source	Confidence Level
Utilities	Medium
Other point sources	Low-Medium
On-road mobile	Low-Medium
Nonroad mobile	Low-Medium
Stationary nonpoint	Low

The above rankings are not specific to benzene, but generally applicable to HAP emission estimates. Because utilities are not an important benzene source, this rating scheme just indicates that point, on-road and nonroad HAP emission estimates are about equally uncertain. The nonpoint source HAP emission estimates are the most uncertain because of the lack of direct measurement of their emissions. While it is acknowledged that there are differences in the methods used to estimate benzene emissions for the different major sectors, no analysis of the disharmony associated with using these different methods is presented here.

Overall, we find the most important sources of uncertainty in benzene emissions in the Houston area to be non-EGU point sources. We provide greater detail on emissions uncertainties by source category below.

Climate change effects on future emissions (in 2020) are not incorporated in this analysis. Long-term average historical temperature and humidity conditions were modeled in the projection years for mobile sources. If actual temperatures in 2020 are higher than long-term averages, then benzene emissions from evaporative sources may be underestimated. Evaporative benzene sources are present in all sectors except electric generating units.

1. Point Sources

Researchers have recently concluded that there is a high level of uncertainty in the HG area point source VOC and toxic air contaminant emission estimates, especially for petrochemical facilities. For example, the second Texas Air Quality study (TexAQS II) in 2006 confirmed that inventories based on standard EPA emission factors significantly underestimated petrochemical emissions. The analysis, which focused mostly on VOC emissions in the HGB ozone nonattainment area, stated that ethane emissions from petrochemical facilities were underestimated in the 2004 TCEQ point source database by one or two orders of magnitude. Measurements made during TexAQS II also suggest that some toxic petrochemical emissions are similarly underestimated. Airborne measurements showed high benzene concentrations in the Houston Ship Channel area—where many petrochemical facilities are located—in contrast to relatively low benzene concentrations in urban areas of Houston. In addition, measurements taken in the Ship Channel from a ship recorded measurements as high as 50 ppb. These high readings suggest that the petrochemical facilities in the ship channel area contribute more benzene than is reported in the area's point source emission inventories.

A number of reasons for the difference between estimated and observed VOC and benzene emissions have been posited by researchers. The City of Houston's application to EPA for review of petrochemical emission factors mentions the following: emission measurements fail to account for emissions generated during start-up, shutdown and malfunctions and increased emissions that result from poor equipment maintenance; emission estimation methods fail to account for environmental variables that significantly affect emissions (such as the effect of wind speed on flare and storage tank emissions), and not accounting for the emissions from delayed coker units at refineries (which are an important benzene emissions source) (City of Houston, 2008).

Based on the above, it is expected that the point source benzene emission estimates for 1990 and 2002 have significant uncertainties. Whether these uncertainties are as great in the 2010 and 2020 forecasted emissions depends on the ability of the highly reactive volatile organic compound (HRVOC) rules that were initiated in the HGB area since 2000 are effective in reducing benzene emissions from fugitive emission sources such as valves and flanges, flares, process vents and cooling towers. The 2004 HGB ozone SIP revision made the assumption that the VOC emissions were higher in the base year than was estimated in the point source emission inventory, but that these emissions would be fully controlled by the attainment year via regulations adopted to control them. If the benzene emissions from point sources in the HGB area are underestimated by a significant amount, then both the with and without CAAA scenario emissions should be higher than estimated in this study. If the HRVOC rules are successful in reducing benzene emission releases, then the incremental effect of the CAAA is underestimated in this study. If the HRVOC rules are unsuccessful in reducing fugitive benzene releases, then the CAAA benefit should be close to that shown in this report.

Industrial leak detection and repair reductions that are part of the Texas SIP for ozone have not been incorporated in the with-CAAA scenario. Because these programs have been adopted in order to reduce fugitive VOC emissions that have not been captured in the VOC emission inventory for the study area, their emission benefit is difficult to model. The 1990 and 2000 emissions for source categories affected by LDAR rules are likely underestimated.

2. Mobile Sources

MOBILE6.2 is the emission model used in this project to estimate mobile source benzene emission rates. The model provides criteria-pollutant (including PM and NH_3) and HAP emission factors for highway motor vehicles such as passenger cars, trucks, and buses.

MOBILE6.2 uses statistical relationships based on about 1800 emission tests in calculating benzene emission factors. The algorithms used to calculate benzene emission factors in MOBILE6.2 were originally developed for the Complex Model. Both exhaust and evaporative emission factors for benzene are calculated as a ratio of benzene to total organic gases (TOG). Thus, any factor in MOBILE6.2 that affects the TOG emission factors will also affect the benzene emission rates. The benzene to TOG ratios for exhaust emission factors from gasolinepowered vehicles and trucks include variables that account for the volume percent of benzene and aromatics in the gasoline. Exhaust benzene to TOG ratios for diesel-fueled vehicles are constants. The benzene emission factors for LDGVs also include a factor that accounts for aggressive driving, as defined by differences in emissions from the Federal Test Procedure driving cycle and the Unified Cycle. These aggressive driving, or off-cycle, adjustments differ for normal HC emitters and high HC emitters. The evaporative benzene to TOG ratios are a function of the weight percent of oxygen in the gasoline, the Reid vapor pressure (RVP) of the gasoline, and the volume percent of benzene in the gasoline.

While the exhaust fractions of benzene to TOG vary by vehicle type and technology group, the data were developed based on a 1990 fleet. Thus, the fractions specific to LEV and Tier 2 vehicles were not known at the time these algorithms were developed. Testing of a small number of LEVs has suggested that the toxic to TOG ratios for LEVS are similar to the ratios used in the Complex Model, but more testing is needed. This leads to some uncertainty in the 2000 benzene emission estimates and a greater degree of uncertainty in the 2010 and 2020 benzene estimates, since by 2020, almost all vehicles in the fleet should meet Tier 2 emission standards.

The TOG emission rates calculated by MOBILE6.2 are based on thousands of emission tests performed on both new and in-use vehicles. In addition to standard testing conditions, many vehicles have been tested at non-standard temperatures, with different types of fuels, including gasoline oxygenate/alcohol blends, and under different driving cycles. Relationships have been developed for vehicles at varying emission control levels, ranging from no control to projections of in-use performance of new technology vehicles.

Even though systematic emission measurements have been performed on the in-use vehicle fleet in the United States, substantial uncertainty remains regarding the applicability of these results. The primary sources of uncertainty are the sensitivity of vehicle emissions to the driving cycle, the wide variety of driving patterns, and the effects of sampling error. Remote sensing surveys indicate that a small fraction of high emitters in the fleet produce a large fraction of total vehicle emissions.

Since MOBILE6's release in January 2001, there have been two studies sponsored to evaluate and validate the model, one sponsored by the Coordinating Research Council (CRC – a cooperative research effort of the American Petroleum Institute and automotive industry in the United States) and EPA, and another sponsored by the American Association of State Highway and Transportation Officials (AASHTO).

The CRC/EPA project (ENVIRON, 2004) compared MOBILE6 HC, CO, and NO_x emission estimates with various real-world data sources, including tunnel studies, ambient pollutant concentration ratios, emission ratios from remote sensing devices, and heavy-duty vehicle emission data based on chassis dynamometer testing. Compared with tunnel studies, the CRC/EPA study found that MOBILE6 over-predicts fleet average emissions, with the over-prediction being most pronounced for CO (and, in particular, newer vehicles). Estimates of NO_x emissions most clearly matched the tunnel data. It should be noted, however, that tunnel study data typically reflect steady-state conditions, whereas MOBILE6 models emission rates in-use across the entire range of operating conditions. Compared with ambient data, the HC/NO_x ratios developed from MOBILE6 appear to be reasonably accurate, and the CRC/EPA data generally supported the HC deterioration rates in MOBILE6.

AASHTO (Sierra, 2004) evaluated several components of MOBILE6 including (1) PM emission factors, (2) toxic air pollutant emission factors, (3) assessment of emission factors when compressed natural gas is the fuel, and (4) methods to estimate CO_2. It was found that MOBILE6 appears to overestimate exhaust PM emissions from newer vehicles. For pre-1990 model years, MOBILE6 predictions fall within the range of recent test program expected values. The AASHTO study also found that MOBILE6 may be underestimating PM_{10} emissions from heavy-duty diesel trucks. The study also found that MOBILE6 brake-wear emission factors likely underestimate brake-wear emissions from the heavier vehicle classes.

Onroad vehicle benzene emission estimates are also subject to the uncertainty associated with the methods used to estimate vehicle activity. For this study, VMT estimates prepared for the Houston area by the Houston-Galveston Area Council (HGAC) and further processed by the Texas Transportation Institute were used. The types of models used by the HGAC are generally considered the most accurate source of onroad activity for an urban area. A recent study (TRB, 2007) used travel forecasts for six metropolitan areas made in 1980 for 2000 metropolitan population, households, and employment versus actual data for 2000. Considerable variation between the 20-year forecasts and the actual situation in 2000 was found, with differences ranging from -23.8 percent to +8.1 percent. These data show the degree of uncertainty associated with such forecasts, regardless of how sophisticated the forecasting process in use may be.

Regarding travel activity estimates, this study has not accounted for the effect of higher motor fuel prices in the with CAAA scenario on vehicle miles traveled. The same VMT data were used in both the “with-CAAA” and the “without-CAAA” scenarios in a given year (i.e., not reflecting any differences in VMT due to the CAAA). The VMT data used in this analysis were based on historical and projection year VMT for the Houston area at the link level, as provided by the Texas Transportation Institute. This data set was believed to provide the best representation of VMT in the modeled area. However, it may be inconsistent with the national DOE VMT projections. Typical metropolitan area travel forecasts, like those being used here, do not include energy prices as a variable.

3. Nonpoint Sources

Compared with other source categories, nonpoint stationary source emissions have the highest uncertainty in emission rates (NARSTO, 2004). Because direct measurement of nonpoint stationary sources is resource intensive, nonpoint stationary source inventories are constructed generally through calculations. Individual sources are smaller and more widely dispersed. In some situations, surrogates for emission and activity factors are used for emission estimates. The quality of the estimate depends on how well the surrogate activity factor correlates with the emission rate for the source.

4. Nonroad Sources

The toxic-to-VOC ratios in EPA's National Mobile Inventory Model (NMIM) for lawn and garden equipment, which makes the single largest contribution of any nonroad sector to the air toxics inventory, is supported by a large amount of test data. The VOC estimates for uncontrolled engines in the NONROAD model are based on a large amount of in-use test data and peer reviewed methodologies. Estimates for controlled engines are based on certification test data and emission standards. However, for a number of source categories – in particular heavy-duty diesel engines and aircraft engines – the toxic to VOC ratios used to develop inventory estimates are based on very limited data. In addition, EPA has limited emissions data for nonroad equipment with emission controls. EPA has been collecting additional measurements to address some of these limitations. There are also significant uncertainties associated with allocating nonroad equipment emissions from the national to the local level. Finally, the relationship between fuel parameters and emission rates for gasoline nonroad equipment is much more poorly understood than the relationship for highway gasoline vehicles (EPA, 2007).

Georgia Tech used expert elicitation methods to quantify NONROAD model emissions uncertainty when applied to estimate criteria pollutant emissions for the State of Georgia (Chi, et al., 2004). The uncertainties in NONROAD emission estimates as calculated in this study were between 24 and 30 percent (standard deviation as the percentage of the mean) for total hydrocarbons. This study estimates nonroad equipment benzene emissions as a percentage of VOCs.

Frey and Bammi (2002) performed an uncertainty analysis for lawn and garden equipment THC emission estimates and found the relative uncertainty to be -32 to 38 percent for 2 stroke engines and -38 to 45 percent for 4 stroke engines.

In a separate analysis, Hanna, et al. (2004) estimated a probabilistic benzene emission inventory developed for the Houston area to have a 95 percent confidence interval of -46 to 108 percent. In this Houston assessment, the authors estimated the key source of uncertainty for Houston benzene emissions to be gasoline onroad mobile sources. This analysis was performed using calendar year 1996 benzene emission estimates for the area and had the 1996 EPA National Toxics Inventory as a data source. So the distribution of benzene emissions among source categories is somewhat different from what was used in this study.

CHAPTER II. STATIONARY SOURCE CATEGORY AND NON-NONROAD MODEL SOURCE CATEGORY EMISSION ACTIVITY INDICATORS

This chapter describes the development of emission activity factors that reflect the projected ratios of 2000, 2010, and 2020 emission activity to 1990 emission activity (for *without-CAAA* case emissions modeling) and ratios of 2010 and 2020 emission activity to 2002 emission activity (for *with-CAAA* case emissions modeling). Emission activity levels for energy producing and consuming source categories are developed from historical/forecast energy production/consumption data. Because it is not feasible to develop estimates of actual emission activity levels for every non-energy related source category, growth factors for these source categories are typically derived from surrogate socioeconomic indicator data that are more readily available than emission activity data.

The following section describes the energy and socioeconomic data that were used to estimate stationary (point and nonpoint) source category emission activity for each year of interest. This discussion also pertains to the use of such data for projecting nonroad source categories that are not incorporated into EPA's NONROAD emissions model (hereafter referred to as "miscellaneous" nonroad source categories).[1] This section is followed by a discussion of alternative data sources and methods that were used to estimate emission activity estimates for a small number of source categories. The final section of this chapter describes how growth indicators were assigned to benzene source categories, and summarizes the backcast/growth factors that were developed for the highest-emitting point, nonpoint, and miscellaneous nonroad benzene source categories in the 1990 (*without-CAAA* case) and 2002 (*with-CAAA* case) base year inventories.

A. ENERGY AND SOCIOECONOMIC DATA EMISSION ACTIVITY INDICATORS

1. Energy Consumption Data

In keeping with past EPA practice, this study relies on energy data from the U.S. Department of Energy (DOE)'s Energy Information Administration (EIA) to backcast/forecast energy consumption and energy production emission source categories. To reflect the 1990 to 2000 trend in energy consumption for source categories, Pechan generally relied on historical time-series energy data for Texas from an EIA energy consumption database (EIA, 2005a). For Crude Oil and Natural Gas Production source categories, Pechan obtained 1990 and 2000 Texas relevant activity data from another EIA source that provided the number of operating oil well days (used for Crude Oil Production) and the number of operating gas well days (used for Natural Gas Production) (EIA, 2005b). For source categories that describe railroad and marine distillate fuel consumption emission processes, Pechan obtained 1990 and 2000 consumption estimates for Texas from an EIA distillate fuel data resource (EIA, 2005c).

---

[1] These "miscellaneous" nonroad categories describe aircraft, marine vessel, and railroad emission processes.

Each year, the EIA produces energy projections for the United States. These projections, which forecast U.S. energy supply, demand, and prices through 2025, are published in an EIA document entitled *Annual Energy Outlook 2005 (AEO 2005)* (EIA, 2005d). For most energy sectors/fuel types, *AEO 2005* reports energy forecasts by Census division. These divisions are defined by State boundaries. When *AEO 2005* produces West South Central (WSC) region forecasts, these regional data were used to project changes in Houston-Galveston area emissions activity (Texas is included in the WSC region). For example, Stage II (Gasoline Vehicle Refueling) emission activity is projected using *AEO 2005* projections of WSC region transportation sector motor gasoline consumption. This study relies on national energy forecasts whenever *AEO 2005* only produces national projections for the energy growth indicator of interest.

2. Socioeconomic Data

Because population growth and the performance of the U.S. economy are two of the main determinants of energy demand, the EIA also prepares socioeconomic projections. These projections feed into energy demand models incorporated into the EIA's National Energy Modeling System (NEMS). NEMS incorporates population projections and economic output forecasts for most industry sectors by Census division. For non-energy intensive economic sectors (e.g., Wholesale Trade), EIA prepares national-level output forecasts. This study relies on *AEO 2005* historical and forecast socioeconomic data as surrogates for emission activity level changes for most non-energy source categories.

County-level population data are one of the key inputs to the BenMap model that will be used to estimate the benefits of air quality changes in the main 812 analysis. Although EPA will be using a benzene-specific health benefits model for the air toxics case study, we start with BenMap population estimates for consistency with the main analysis. In order to be consistent with *AEO 2005* population data, the BenMap population estimates, which were prepared by Woods & Poole Economics Inc. (Woods & Poole, 2001), were revised to make them consistent with the *AEO 2005* regional population estimates. Table II-1 displays comparisons of the *AEO 2005* WSC region population estimates with analogous data compiled from BenMap's county population estimates (Price, 2005). This table indicates that the *AEO 2005* population estimates reflect slightly higher growth rates in the WSC region than the BenMap population estimates.

The first step in developing *AEO 2005* normalized county population projections was to compute factors from the population data in BenMap. These factors represent year-specific ratios of each Houston-Galveston area county's BenMap population to the WSC region's BenMap population. Next, the *AEO 2005* WSC population data were multiplied by these factors to yield this study's population projections for the Houston-Galveston area counties. These estimates are used both in estimating emission activity levels for certain benzene source categories and in health benefits modeling.

Table II-1. Comparison of AEO 2005 and BenMAP Population Estimates West South Central Region

Year	Source	Population	Growth % Relative to 1990 (*without-CAAA* Case)	Growth % Relative to 2002 (*with-CAAA* Case)
1990	*AEO 2005*	26,888,101		
	BenMAP	26,702,788		
2000	*AEO 2005*	31,619,909	17.6	
	BenMAP	31,444,848	17.8	
2002	*AEO 2005*	32,479,468	n/a	
	BenMAP	32,181,108	n/a	
2010	*AEO 2005*	35,900,165	33.5	10.5
	BenMAP	35,162,938	31.7	9.3
2020	*AEO 2005*	40,239,254	49.7	23.9
	BenMAP	39,011,808	46.1	21.1

n/a – not applicable

B. ALTERNATIVE EMISSION ACTIVITY INDICATORS

In some instances, energy and socioeconomic forecasts are not expected to be valid surrogates of emission activity changes. In preparing recent projections to support an analysis of the Clean Air Interstate Rule (CAIR), EPA chose to use alternative emission activity growth surrogates for certain source categories (Houyoux, 2004). Pechan first reviewed the data sources/approaches that were used to support the CAIR projections for application in this study. In addition, for a companion national Section 812 study, Pechan performed new research into the availability of alternative forecast data sources for the highest criteria pollutant-emitting source categories in 2002 (Pechan, 2005a). For the most part, Pechan utilized consistent growth factor development methods to these previous studies in developing non-*AEO 2005* emission activity growth indicators. Table II-2 summarizes the non-*AEO 2005* growth indicators that were applied in this benzene study. The following describes how non-*AEO 2005* emission activity indicator data were developed for the years of interest.

1. Residential Wood Fireplaces and Wood Stoves

Pechan estimated emission activity levels for residential wood burning fireplaces and wood stoves using a combination of DOE national historical residential wood consumption estimates, *AEO 2005* WSC region energy projections, the estimated proportion of consumption by type of unit in each analysis year, and an assumed 2 percent annual turnover to lower-emitting combustion units.

a. Energy Consumption Data

WSC region residential renewable energy consumption estimates were obtained from *AEO 2005* for 2002, 2010, and 2020 (wood accounts for the vast majority of residential renewable energy

Table II-2. Emission Activity Growth Factors Derived from Non-AEO 2005 Forecast Data

Source Classification Code (SCC)	SCC Description	Data Source	Geography	Forecast Data Set(s)
2104008002	Residential Wood Fireplaces: Fireplaces: Insert; Non-EPA Certified	Residential wood consumption index derived from EIA data, extrapolation of unit type wood consumption shares, and 2% annual turnover to EPA certified units	Region (Census division)	Residential renewable energy consumption and forecast distribution of wood consumption by unit type
2104008004	Residential Wood Fireplaces: Fireplaces: Insert; EPA Certified; Catalytic	Residential wood consumption index derived from EIA data, extrapolation of unit type wood consumption shares, and 2% annual turnover to EPA certified units	Region (Census division)	Residential renewable energy consumption and forecast distribution of wood consumption by unit type
2104008010	Residential Wood Stoves: General	Residential wood consumption index derived from EIA data, extrapolation of unit type wood consumption shares, and 2% annual turnover to EPA certified units	Region (Census division)	Residential renewable energy consumption and forecast distribution of wood consumption by unit type
2104008030	Residential Wood Stoves: Catalytic Woodstoves: General	Residential wood consumption index derived from EIA data, extrapolation of unit type wood consumption shares, and 2% annual turnover to EPA certified units	Region (Census division)	Residential renewable energy consumption and forecast distribution of wood consumption by unit type
2275020000	Commercial Aircraft	Federal Aviation Administration forecasts of landing and take-off (LTO) operations (FAA, 2005)	State	Air carrier operations
2275050000	General Aviation		State	General aviation operations
2275060000	Air Taxi		State	Air taxi operations
2810001000	Forest Wildfires	Houyoux, 2004	Not applicable	No change
2810015000	Prescribed Burning for Forest Management	Houyoux, 2004	Not applicable	No change

consumption). Because State-level residential wood consumption estimates appeared suspect, Pechan used DOE national 1990, 2000, and 2002 residential wood consumption data to estimate the trend in WSC region residential wood consumption over this period. Pechan then combined the two sets of estimates to develop estimates of WSC residential renewable energy consumption for each year of interest. Figure II-1 displays the estimated trend in residential wood consumption relative to 2002 consumption levels. This figure shows that there was a dramatic drop in residential wood burning in the 1990 to 2002 period. This activity reduction, which may be the result of the cost of wood versus natural gas as a home heating fuel, is the most important factor in reducing residential wood combustion emissions during this period.

b. Estimates of Residential Wood Consumption Proportions by Unit Type

From the U.S. Bureau of the Census's *Census of Housing*, Pechan obtained the 1997, 1999, 2001, and 2003 national number of homes with wood stoves and number of homes with fireplaces with inserts (BOC, 2004).[2] For fireplaces without inserts, Pechan compiled *Census of Housing* data reflecting the number of homes that use fireplaces without inserts as the main heating source and the number of homes that use fireplaces without inserts as a supplementary heating source. Pechan then adjusted the *Census of Housing* data to reflect the estimated number of units per home − 1.1 for fireplaces with inserts; 1.17 for fireplaces without inserts, and 1.09 for stoves (Pechan, 2006). Next, Pechan multiplied the product of these numbers by estimated annual wood consumption per unit − 1.533 cords per unit for wood stoves and fireplaces with inserts; 0.656 cords per unit for fireplaces without inserts used as the main heating source; and 0.069 cords per unit for fireplaces without inserts used for other heating (Pechan, 2006). Pechan then summed the main heating and other heating estimates for fireplaces without inserts to yield total wood consumption for fireplace without inserts. These calculations resulted in estimated national wood consumption for 1997, 1999, 2001, and 2003 for woodstoves, fireplaces with inserts, and fireplaces without inserts.

Next, Pechan computed the proportions of total residential wood consumption by each unit type for the available years. Pechan then interpolated these proportions for the intervening years. Because pre-1997 *Census of Housing* values appeared anomalous, Pechan chose to use the 1997 residential wood consumption proportions to represent the 1990 proportions. Next, Pechan projected the wood consumption shares by residential wood combustion (RWC) unit type in 2010 and 2020 by extrapolating from the 1997-2003 values. Figure II-2 presents the estimated proportions of total consumption by unit type over the 1997-2020 period.

c. SCC-Level RWC Activity Consumption Forecasts

Two additional steps were used to develop source classification code (SCC)-level RWC activity forecasts. First, Pechan estimated wood consumption by RWC unit type for the period 2002-2020 by multiplying *AEO 2005* WSC region 2002-2020 renewable energy consumption by 2002-2010 RWC unit type wood consumption proportions. Next, for SCCs that disaggregate the broad RWC unit types reported in the *Census of Housing* data (i.e., woodstoves, fireplaces with inserts, and fireplaces without inserts), Pechan allocated the consumption estimates to these more detailed SCCs. For the woodstoves and fireplaces with inserts categories, this step involved

---

[2] Pre-1997 data were not compiled because these data reflected an anomalous disconnect in the data series trend.

Figure II-1. Total Residential Wood Consumption Relative to 2002 Consumption

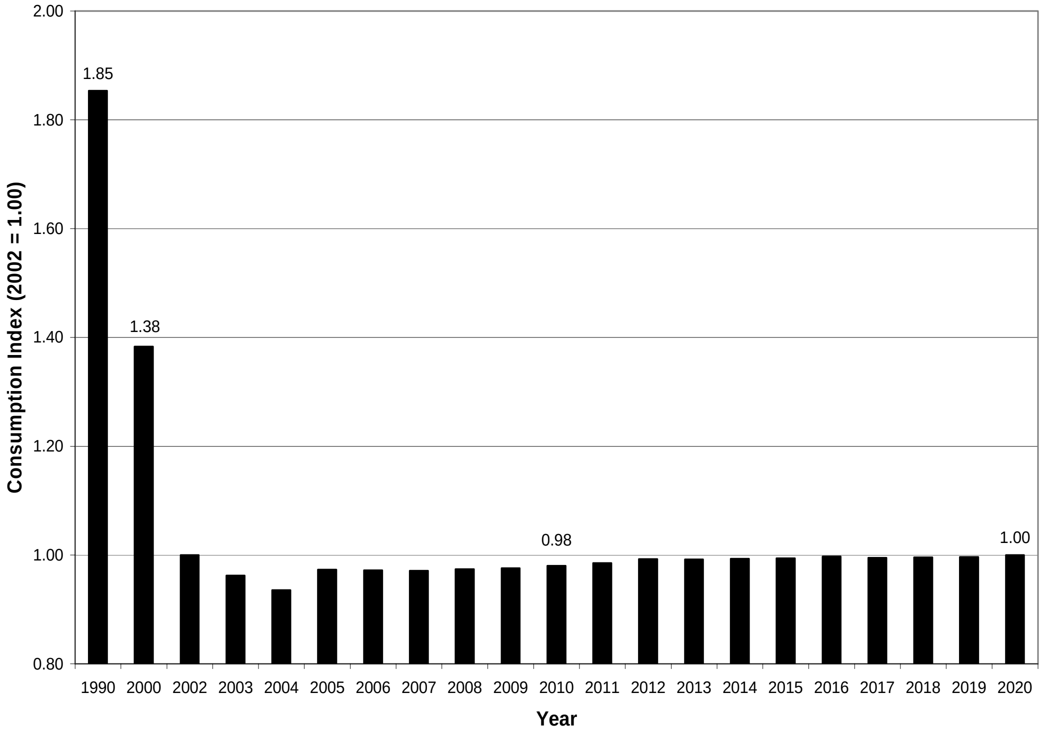
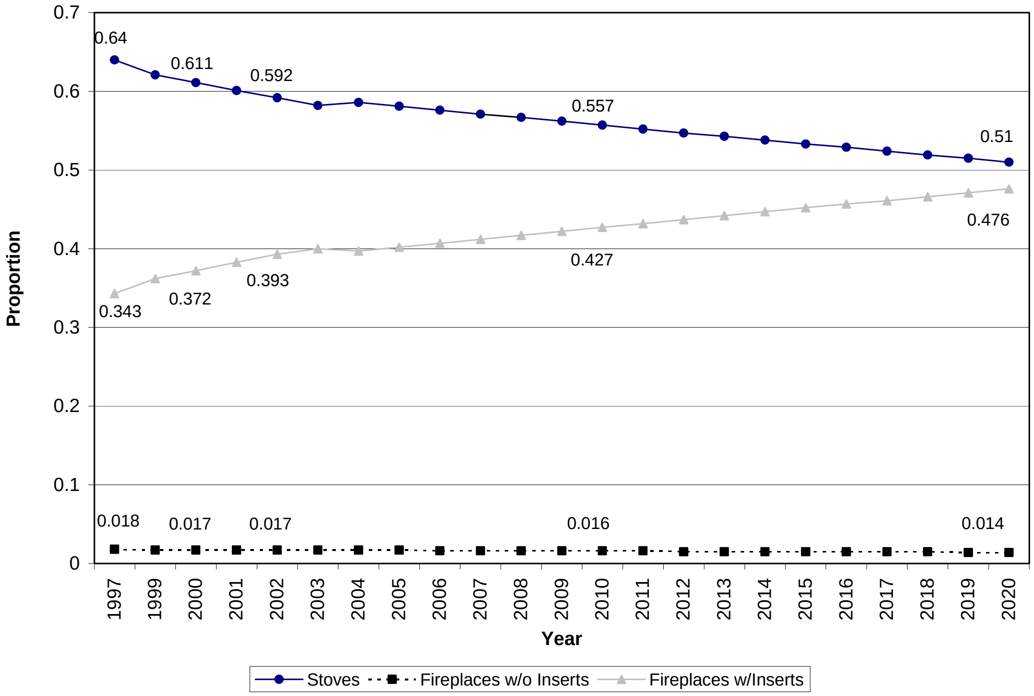


Figure II-2. Proportion of Total Residential Wood Consumption by Type of Unit



allocating consumption into three SCC-specific unit types representing non-EPA certified, EPA certified catalytic, and EPA certified non-catalytic units.

This 2002 year allocation was accomplished by multiplying the broad unit-level consumption estimates by the proportions of total RWC attributed to each SCC as reported in the 2002 NEI: 92 percent for non-EPA certified units; 5.7 percent for EPA certified non-catalytic units; and 2.3 percent for EPA certified catalytic units (Pechan, 2006). To reflect a projected increase in EPA-certified units resulting from EPA's wood heater New Source Performance Standard (NSPS), forecast year proportions were calculated by adjusting the 2002 year proportions using an annual 2 percent RWC unit turnover rate computed from 1992-2005 data (Broderick and Houck, 2005). This adjustment accounts for non-EPA certified units being replaced by NSPS compliant EPA-certified units. Therefore, by the year 2020, it is assumed that 64.4 percent of residential wood consumption in woodstoves and fireplaces with inserts will occur in non-EPA certified units, 25.4 percent in EPA certified non-catalytic units, and 10.2 percent in EPA certified catalytic units.

To ensure that 1990 and 2000 emissions are calculated in a consistent manner with the forecast year values, Pechan developed 1990 and 2000 estimates as follows. Pechan first calculated ratios representing 1990 and 2000 residential wood consumption relative to 2002 consumption (1.85 and 1.38, respectively), and then multiplied these ratios by 2002 year WSC region residential renewable energy consumption. Next, Pechan applied values representing the estimated 1990 and 2000 year proportions of total residential wood consumption attributable to each of the following unit types: woodstoves, fireplaces with inserts, and fireplaces without inserts (see Figure II-2).[3]

To calculate the 1990 SCC-level backcast factors relative to 2002, Pechan allocated the general unit-level consumption estimates to individual SCCs. This step assumed that zero residential wood consumption would occur in 1990 in EPA certified units because 1992 was the first year of certification (Broderick and Houck, 2005). Finally, Pechan calculated the ratio of estimated 1990 year consumption to estimated 2002 consumption for the SCCs that appear in the 2002 base year inventory.

To calculate the 2000 SCC-level backcast factors relative to 2002, Pechan utilized the aforementioned annual 2 percent turnover rate and the 2002 NEI wood consumption proportions to estimate the following proportions in 2000: 95.68 percent for non-EPA certified units; 3.08 percent for EPA certified non-catalytic units; and 1.24 percent for EPA certified catalytic units. These estimates were used to calculate SCC-level wood consumption. Finally, Pechan computed the ratio of 2000 consumption to 2002 consumption for all base year inventory SCCs.

Table II-3 displays the 1990 and 2000 backcast factors (*without-CAAA* case) and the 2010 and 2020 growth factors (*with-CAAA* case) for the SCCs included in the 2002 Houston-Galveston base year benzene inventory. These factors are expressed relative to 2002 consumption levels.

---

[3] As noted earlier, Pechan used the 1997 proportions to represent 1990 proportions.

Table II-3. Residential Wood Consumption Backcast/Growth Factors

SCC	SCC/Unit Type	1990 Factor	2000 Factor	2010 Factor	2020 Factor
2104008002	Non-certified Fireplaces with Inserts	1.766	1.368	0.910	0.851
2104008004	Certified Catalytic Fireplaces with Inserts	0	0.711	2.917	5.412
2104008010	Woodstoves, General	2.006	1.428	0.923	0.862
2104008030	Certified Catalytic Woodstoves	0	0.771	2.515	3.836

2. Aircraft

State-level projections of the number of operations (arrival and departures) by type of aircraft (commercial, air taxi, and general aviation) were obtained from the Federal Aviation Administration (FAA)'s Terminal Area Forecasts (FAA, 2005). The FAA's itinerant and local operations data were summed to develop total operations by aircraft type. Because number of landing and take-offs (LTOs) is the emission activity for these source categories, and because an LTO is equivalent to two total operations (i.e., one arrival and one departure), Pechan divided the number of total operations by 2 to yield the number of LTOs. Growth factors were developed for each type of aircraft by dividing the forecast year LTO projections by 2002 LTO estimates. To ensure that 1990 emission values are calculated for the same SCCs and on a consistent basis with the base year and forecast year values, Pechan replaced the 1990 base year aircraft emission estimates with estimates computed by multiplying the 2002 emissions by the ratio of 1990 LTOs to 2002 LTOs. Similarly, Pechan computed 2000 *without-CAAA* emission estimates by multiplying 2002 emissions by the ratio of 2000 LTOs to 2002 LTOs. Table II-4 presents the LTO-based backcast/growth factors for the aircraft source categories reported in the 2002 benzene emission inventory.

Table II-4. Aircraft Growth Factors Relative to 2002 Base Year

Source Category	1990	2000	2010	2020
Commercial Aircraft	0.961	1.159	1.067	1.222
General Aviation	0.993	1.003	1.023	1.154
Air Taxi	0.684	1.000	1.419	1.880

3. Forest Wildfires

This source category's emissions activity is largely a function of meteorological conditions and unintentional activities. In keeping with analyses performed in support of the CAIR, Pechan assumed no change in 2002 year forest wildfire activity levels throughout this study's analysis period (Houyoux, 2004).

4. Prescribed Burning for Forest Management

For consistency with emission forecasts prepared to support EPA's CAIR, Pechan assumed no change in 2002 year prescribed burning activity levels throughout this study's analysis period (Houyoux, 2004).

C. ASSIGNMENT OF GROWTH INDICATORS TO BASE YEAR EMISSION SOURCES

The following subsections describe the methods that were used to assign growth indicators to energy and non-energy related emission source categories.

1. Assignments for Energy Related Source Categories

Because *AEO 2005* and historical EIA publications provide detailed energy production/ consumption data by sector and fuel type, energy-related source categories can be easily matched with an appropriate EIA growth indicator. Pechan assigned growth indicators to energy production/consumption emission source categories using recent MACT and SCC growth indicator crosswalks developed in support of Version 5.0 of the Economic Growth Analysis System (EGAS) (Pechan, 2005b; 2005c).[4]

2. Assignments for Non-Energy Related Source Categories

For non-energy related emission source categories, Pechan generally utilized *AEO 2005* sector output data as surrogates for changes in emission activity.[5,6] The EGAS 5.0 sector output-based crosswalks utilize Standard Industrial Classification (SIC) code-based output projections from Regional Economic Models, Inc. (REMI) as growth indicators for non-energy related MACT and SCC codes. For consistency with the *AEO 2005* energy projections, Pechan used *AEO 2005* economic output projections in this study rather than the REMI economic output projections. The EGAS 5.0 output forecasts are available for approximately 165 separate economic sectors, while the *AEO 2005* output projections are available for about 50 economic sectors (see Table II-5 for list of *AEO 2005* sectors). Therefore, the growth indicators used in this study are less sector-specific than the growth indicators used in EGAS 5.0 or in the recent CAIR projections. However, the *AEO 2005* historical/forecast economic output data are used to ensure consistency with the economic projections used in forecasting *AEO 2005* energy production/consumption. The following subsections describe how the *AEO 2005* socioeconomic data were assigned as growth indicator surrogates for non-energy related source categories.

a. MACT Code Assignments

As part of the regulatory development process, EPA has identified the economic sectors affected by MACT standards. EPA regulatory documents generally list the North American Industrial Classification (NAICS) codes potentially affected by MACT standards. Because this information can be used to specifically relate MACT codes to NAICS codes, Pechan used MACT codes to link with the appropriate *AEO 2005* output sector whenever a valid MACT code was reported in the base year inventory. Before the transition from SIC codes to NAICS codes, EPA regulatory documents listed the SIC codes affected by MACT standards. For these regulations, Pechan used a U.S. Bureau of the Census crosswalk that links SIC codes to NAICS

---

[4] These crosswalks utilize *AEO 2004* data, which are reported for essentially the same sectors/fuel types as the *AEO 2005* projections data.

[5] Note that unlike energy production/consumption data, historical sector output data were available from *AEO 2005.*

[6] In addition to sector output, population is used as the growth indicator for some non-energy source categories.

codes to assign the appropriate *AEO 2005* NAICS-based growth indicator(s) to MACT codes (BOC, 2005).

b. SCC Assignments

When a valid MACT code was not available for an emission record in the inventory, Pechan assigned the growth indicator based on the SCC. Pechan used a combination of the EGAS 5.0 SCC-based crosswalk and the U.S. Bureau of the Census' SIC code to NAICS code crosswalk to assign *AEO 2005* NAICS-based growth indicators to SCCs. Because the EGAS 5.0 crosswalk links REMI SIC code-based economic sectors to SCCs, Pechan used the Census's SIC code to NAICS code crosswalk to identify the *AEO 2005* sector indicator(s) to apply for a given non-energy related SCC (note that, in keeping with EGAS 5.0, population is used as the growth indicator for many such SCCs).

Table II-6 presents *without-CAAA* case growth factors for the highest benzene-emitting stationary source and miscellaneous nonroad source categories (50 tpy or more) in 1990. These growth factors are reported relative to 1990 activity levels (i.e., 1990 = 1.0). As indicated in the table, growth indicators were assigned to most of the top benzene-emitting stationary/ miscellaneous nonroad source categories by linking with the emission record's MACT code.

Table II-7 displays *with-CAAA* case growth factors for the highest benzene-emitting stationary source and miscellaneous nonroad source categories (20 tpy or more) in 2002. The Table II-7 growth factors are stated relative to 2002 base year activity levels. It is important to note that the two largest 1990 benzene-emitting SCCs do not appear in this table because they are no longer top-emitting categories in the 2002 inventory. This table also indicates that many more *with-CAAA* case growth indicator assignments were based on an emission record's SCC than in the *without-CAAA* case.

Table II-5. AEO 2005 Economic Sectors

Geography	Sector	NAICS Code(s)
Regional	MFGO1 Food Products	311
	MFGO2 Beverage and Tobacco Products	312
	MFGO3 Textile Mills & Textile Products	313,314
	MFGO4 Apparel	315
	MFGO5 Wood Products	321
	MFGO6 Furniture and Related Products	337
	MFGO7 Paper Products	322
	MFGO8 Printing	323
	MFGO9 Basic Inorganic Chemicals	32511,32519
	MFGO10 Basic Organic Chemicals	32512 – 32518
	MFGO11 Plastic and Synthetic Rubber Materials	3252
	MFGO12 Agricultural Chemicals	3253
	MFGO13 Other Chemical Products	3254 – 3259
	MFGO14 Petroleum Refineries	32411
	MFGO15 Other Petroleum and Coal Products	32412,32419
	MFGO16 Plastics and Rubber Products	326
	MFGO17 Leather and Allied Products	316
	MFGO18 Glass & Glass Products	3272
	MFGO19 Cement Manufacturing	32731
	MFGO20 Other Nonmetallic Mineral Products	327 less 3272 & 32731
	MFGO21 Iron & Steel Mills, Ferroalloy & Steel Products	3311,3312
	MFGO22 Alumina & Aluminum Products	3313
	MFGO23 Other Primary Metals	3314,3315
	MFGO24 Fabricated Metal Products	332
	MFGO25 Machinery	333
	MFGO26 Other Electronic & Electric Products	334 less 3345 & 335
	MFGO27 Transportation Equipment	336
	MFGO28 Measuring & Control Instruments	3345
	MFGO29 Miscellaneous Manufacturing	339
	MFGO30 Crop Production	111
	MFGO31 Other Agriculture, Forestry, Fishing & Hunting	112 – 115
	MFGO32 Coal Mining	2121
	MFGO33 Oil & Gas Extraction & Support Activities	211,213
	MFGO34 Other Mining & Quarrying	2122,2123
	MFGO35 Construction	23
	Sum of All Chemicals	325
	Sum of All Petroleum	324
	Sum of All Stone, Clay, Glass and Cement	327
	Sum of All Primary Metals	331
	Total Manufacturing Output	31 – 33
	Total Industrial Output	11,21,23,31 – 33
National	NMFGO1 Transportation & Warehousing	48,49
	NMFGO2 Broadcasting & Telecommunications	513
	NMFGO3 Electric Power Generation & Distribution	2211
	NMFGO4 Natural Gas Distribution	2212
	NMFGO5 Water, Sewage & Related System	2213
	NMFGO6 Wholesale Trade	42
	NMFGO7 Retail Trade	44,45
	NMFGO8 Finance & Insurance, Real Estate	52,53
	NMFGO9 Other Services	51,54 – 81
	NMFGO10 Public Administration	921,922,923
	Total NonManufacturing/Service Gross Output	22,42,44,45,48,49,51–81, 92
	Total Gross Output	All

Table II-6. Summary of *Without-CAAA* Case Growth Factors for Highest Emitting Stationary Source and Miscellaneous Nonroad Source Categories in 1990

MACT Code	MACT Code Description	SCC	SCC Description	1990 Benzene Emissions (tpy)	% of Total	Growth Indicator	2000 Growth Factor	2010 Growth Factor	2020 Growth Factor
1501	Synthetic Organic Chemical Manufacturing (HON)	2301040000	Fugitive Emissions from Synthetic Organic Chemical Manufacturing, Total	2,371.43	44	Basic Organic Chemicals	1.202	1.175	1.347
503	Petroleum Refineries - Other Sources Not Distinctly Listed	2306000000	Petroleum Refining, All Process	803.41	15	Refined Petroleum Products Supplied, Total	1.275	1.443	1.652
1501	Synthetic Organic Chemical Manufacturing (HON)	30188801	Chemical Manufacturing, Fugitive Emissions, Specify in Comments	265.88	5	Basic Organic Chemicals	1.202	1.175	1.347
603	Marine Vessel Loading Operations	40600298	Transportation and Marketing of Petroleum Products, Marine Vessels, Not Classified	133.65	2	Transportation & Warehousing	1.364	1.592	2.053
1501	Synthetic Organic Chemical Manufacturing (HON)	30116901	Chemical Manufacturing, Ethyl Benzene	105.21	2	Basic Organic Chemicals	1.202	1.175	1.347
1641	Miscellaneous Organic Chemical Manufacturing	30101861	Chemical Manufacturing, Plastics Production, Purification System (Polyethylene)	91.82	2	Basic Organic Chemicals	1.202	1.175	1.347
		40799997	Petroleum and Solvent Evaporation, Organic Chemical Storage, Miscellaneous	79.22	1	Basic Organic Chemicals	1.202	1.175	1.347
		30700104	Pulp and Paper and Wood Products, Sulfate (Kraft) Pulping, Recovery Furnace/Direct Contact Evaporator	72.60	1	Paper Products	1.130	1.483	1.883
		30114005	Chemical Manufacturing, Acetylene Production	61.72	1	Basic Organic Chemicals	1.202	1.175	1.347
1501	Synthetic Organic Chemical Manufacturing (HON)	38500101	Cooling Tower, Process Cooling, Mechanical Draft	61.08	1	Basic Organic Chemicals	1.202	1.175	1.347
601	Gasoline Distribution (Stage I)	2501050120	Petroleum and Petroleum Product Storage, Bulk Terminals, Breathing Loss, Gasoline	59.21	1	Transportation, Motor Gasoline	1.243	1.531	1.840
1407	Hydrochloric Acid Production	30101199	Chemical Manufacturing, Hydrochloric Acid, Other Not Classified	57.00	1	Basic Organic Chemicals	1.202	1.175	1.347
1501	Synthetic Organic Chemical Manufacturing (HON)	40703602	Organic Chemical Storage, Fixed Roof Tanks-Aromatics, Benzene	53.57	1	Basic Organic Chemicals	1.202	1.175	1.347
			Subtotal	**4,215.81**	**78**				

For consistency with the national Section 812 study, *without-CAAA* case emissions were computed for the following SCCs from 2002 rather than 1990 emissions: 2280002100 (Marine Vessels, Diesel, Port Emissions); 2280002200 (Marine Vessels, Diesel, Underway Emissions); 2285002006 (Line Haul Locomotives; Class I Operations); 2285002007 (Line Haul Locomotives: Class II/III Operations); 2285002010 (Yard Locomotives). The backcast/growth factors for these categories, which are calculated relative to 2002 activity levels, are as follows: SCCs 2280002100 and 2280002200: 1990 = 0.846, 2000 = 0.965, 2010 = 1.112, and 2020 = 1.185; and SCCs 2285002006, 2285002007, and 2285002010: 1990 = 0.689, 2000 = 0.793, 2010 = 1.095, and 2020 = 1.151. See Section B for discussion of additional categories for which *without-CAAA* case emissions were computed from 2002 emissions.

Table II-7. Summary of *With-CAAA* Case Growth Factors for Highest Emitting Stationary Source and Miscellaneous Nonroad Source Categories in 2002

MACT Code	MACT Code Description	SCC	SCC Description	2002 Benzene Emissions (tpy)	% of Total	Growth Indicator	2010 Growth Factor	2020 Growth Factor
1409	Hydrogen Fluoride Production	30180001	Chemical Manufacturing; General Processes; Fugitive Leaks	86.75	7	Basic Organic Chemicals Output	1.175	1.347
		2505020000	Petroleum and Petroleum Product Transport; Marine Vessel; Total: All Products	81.84	7	Total Refined Petroleum Products Supplied	1.166	1.335
0602	Organic Liquids Distribution (Non-gasoline)	40721205	Organic Chemical Storage; Floating Roof Tanks – Glycol Ethers; Carbitol: Standing Loss	51.03	4	Sum of Output for Basic Organic Chemicals and Petroleum Refineries	1.215	1.437
		2501050120	Petroleum and Petroleum Product Storage; Bulk Stations/Terminals: Breathing Loss, Gasoline	42.96	3	Transportation Sector; Gasoline Consumption	1.176	1.413
		2505040120	Petroleum and Petroleum Product Transport; Pipeline; Gasoline	39.71	3	Transportation Sector; Gasoline Consumption	1.176	1.413
1667	Chemical Preparations	30199999	Chemical Manufacturing: Other Not Classified	39.07	3	Other Chemical Products Output	1.294	1.646
0502	Petroleum Refineries – Catalytic Cracking; Catalytic Reforming & Sulfur Plant Units	30688801	Petroleum Industry; Fugitive Emissions	38.49	3	Total Refined Petroleum Products Supplied	1.166	1.335
		2280002100	Marine Vessels, Commercial; Diesel; Port Emissions	38.18	3	Transportation Sector; Marine - Domestic Shipping; Distillate Fuel Consumption	1.112	1.185
		2310001000	Oil and Gas Production: SIC 13; All Processes: On-Shore; Total	36.47	3	Onshore Crude Oil and Natural Gas Production	0.946	0.857
0601	Gasoline Distribution (Stage I)	2501080050	Petroleum and Petroleum Product Storage; Airports: Aviation Gasoline; Stage I: Total	35.13	3	Transportation Sector; Gasoline Consumption	1.176	1.413
		30188801	Chemical Manufacturing; Fugitive Emissions	31.65	3	Basic Inorganic Chemicals Output	1.105	1.190
		2810015000	Prescribed Burning for Forest Management	27.72	2	No growth assumption	1.000	1.000
1641	Miscellaneous Organic Chemical Manufacturing	30103402	Chemical Manufacturing; Aniline/Ethanolamines; General: Aniline	25.55	2	Basic Organic Chemicals Output	1.175	1.347
		2104008002	Residential Wood Combustion; Fireplaces: Insert; non-EPA Certified	23.69	2	Residential renewable energy consumption, unit type wood consumption shares, & 2% annual turnover to EPA certified units	0.910	0.851
			Subtotal	**598.24**	**49**			

[This page intentionally left blank.]

CHAPTER III. POINT SOURCE ANALYSIS

This chapter describes the benzene emissions analysis performed for point sources in the Houston-Galveston study area. EPA's 1990 NEI for HAPs point source file, which was recently redeveloped by EPA, was used to estimate 1990 benzene emissions in the study area, and was used as the base year file for estimating *without-CAAA* scenario emissions for 2000, 2010, and 2020. The EPA 2002 NEI was used to estimate point source benzene emissions for the study area for the 2000 *with-CAAA* scenario, and as the base year emissions file for preparing 2010 and 2020 *with-CAAA* scenario emission estimates. The remainder of this chapter describes how the point source analysis was performed and summarizes the study area benzene emission estimates for all of the scenario years.

A. METHODS

1. Base Year Emission Estimates

a. With-CAAA Projections

The original analytic plan proposed use of the 1999 EPA NEI as the *with-CAAA* base year point source data source. The SAB suggested the use of the Texas AQS 2000 study emission estimates instead. Because of the time that has elapsed since the 1999 EPA NEI was proposed as the primary point source emissions data source in the analytic plan and study initiation, the draft 2002 NEI became available. In the end, it was decided to use the draft 2002 NEI as the base year point source database for the *with-CAAA* analysis (ERG, 2005). The NEI covers all criteria air pollutants and HAPs for the United States. While the Texas AQS study database received strong consideration as the base year data set, the Texas AQS study provided hourly emission estimates for an August-September 2000 Houston area modeling episode. While this database would provide detailed information about the emissions during this period, it would require adjustments to allow it to be used to represent emission conditions during other time periods. In addition, as a modeling database, it did not contain some of the control device information that assists in making emission forecasts to future years. Furthermore, the Texas Commission on Environmental Quality (TCEQ) submittal for the draft 2002 NEI was improved from previous years for many compounds because of the TCEQ work on species profile improvements. These species profile improvements occurred as part of the recent efforts in the Houston-Galveston area to develop improved point source emission estimates for ozone modeling. Individual point sources were surveyed in order to develop the improved compound-specific emission estimates.

The original analytical plan for this study also recommended that the study area be limited to Harris County. This planned approach was based on the limited availability of historical benzene emission estimates for counties other than Harris. The SAB recommended that a larger study area be used. In order to establish an appropriate study area boundary, the point source benzene emission estimates for the entire study area were reviewed based on the 1999 NEI, and it was found that 99 percent of the Houston-Galveston area's point source benzene emissions would be captured if the study area included Brazoria, Galveston, and Harris counties. Therefore, it was recommended that the study area include these three counties. Table III-1 summarizes county-level point source benzene emissions for the counties in the Houston-Galveston ozone nonattainment area.

Table III-1. County Summary of the 1999 NEI Benzene Emissions for Major Point Sources in the Houston-Galveston Area

County Name	1999 Benzene Emissions (tpy)	Percentage of Total
Brazoria County	115.70	**16.41**
Chambers County	0.19	0.03
Fort Bend County	6.39	0.91
Galveston County	148.74	**21.10**
Harris County	433.06	**61.44**
Liberty County	0.80	0.11
Montgomery County	0.00	0.00
Waller County	0.00	0.00
Total	704.87	**100.00**

i. 2002 Draft NEI Point Source

Table III-2 provides a by source category summary of the 2002 point source benzene emissions for the study area. Five source categories contribute nearly 80 percent of the emissions:

- Chemical & Allied Product Mfg - Organic Chemical Mfg (26.94 percent);
- Petroleum & Related Industries - Petroleum Refineries & Related Industries (14.91 percent);
- Storage & Transport - Organic Chemical Storage (13.12 percent);
- Storage & Transport - Petroleum & Petroleum Product Storage (12.60 percent); and
- Chemical & Allied Product Mfg - Other Chemical Mfg (11.50 percent).

The sources of data in the NEI for benzene in the Houston-Galveston area are as follows:

- State data,
- Industry data;
- Data gathered by EPA's Emission Standards Division while developing MACT standards;
- Toxics Release Inventory (TRI) data; and
- Electricity generating unit (EGU) data developed from information by DOE and EPA's Clean Air Markets Division (CAMD).

Compilation of the 2002 NEI for HAPs by EPA required many steps, including the following key processing activities:

Table III-2. Tier 2 Summary of 2002 NEI Draft Point Source Benzene Emissions

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	2002 Benzene Emissions (tpy)	% of Total
01	FUEL COMB. ELEC. UTIL.	03	Gas	11.09	1.43
01	FUEL COMB. ELEC. UTIL.	04	Other	0.33	0.04
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.42	0.05
02	FUEL COMB. INDUSTRIAL	02	Oil	1.20	0.16
02	FUEL COMB. INDUSTRIAL	03	Gas	32.45	4.20
02	FUEL COMB. INDUSTRIAL	04	Other	11.33	1.47
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	0.63	0.08
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.00	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	208.23	26.94
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	1.99	0.26
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	0.68	0.09
04	CHEMICAL & ALLIED PRODUCT MFG	04	Agricultural Chemical Mfg	0.00	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	88.85	11.50
05	METALS PROCESSING	01	Non-Ferrous Metals Processing	0.00	0.00
05	METALS PROCESSING	02	Ferrous Metals Processing	2.06	0.27
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	10.82	1.40
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	115.27	14.91
06	PETROLEUM & RELATED INDUSTRIES	03	Asphalt Manufacturing	0.28	0.04
07	OTHER INDUSTRIAL PROCESSES	04	Rubber & Miscellaneous Plastic Products	1.56	0.20
07	OTHER INDUSTRIAL PROCESSES	05	Mineral Products	0.15	0.02
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	1.92	0.25
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	18.89	2.44
08	SOLVENT UTILIZATION	03	Dry Cleaning	0.03	0.00
08	SOLVENT UTILIZATION	04	Surface Coating	0.13	0.02
08	SOLVENT UTILIZATION	05	Other Industrial	0.02	0.00
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	21.96	2.84
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	97.41	12.60
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	18.48	2.39
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	101.44	13.12
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	4.83	0.62
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	3.42	0.44
10	WASTE DISPOSAL & RECYCLING	03	Publicly-Owned Treatment Works (POTW)	7.17	0.93
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	9.55	1.24
10	WASTE DISPOSAL & RECYCLING	05	TSDF	0.01	0.00
10	WASTE DISPOSAL & RECYCLING	06	Landfills	0.19	0.02
10	WASTE DISPOSAL & RECYCLING	07	Other	0.09	0.01
			Total	**772.88**	**100.00**

- Submittal of 2002 HAP inventory data by reporting agencies;
- Blending/merging data from multiple data sources;
- Augmentation of blended data for missing data elements;
- Quality control/quality assurance of the data;
- Preparation of draft NEI for HAPs for external and internal review
- Incorporation of external and internal review comments; and
- Preparation of final NEI for HAPs

Because facilities and some source categories were missing from State and local agency-supplied data, EPA merged data from other sources to compile the 2002 NEI for HAPs. An automated quality control tool was used to check each State and local agency, MACT source category, and TRI file. Then, for the data blending/merging process, the TCEQ-supplied files were compared with MACT source category files and the TRI database. MACT codes were assigned at either the site or the process level based on:

- emissions data provided by the MACT engineer;
- a facility list provided by the MACT engineer; or
- the SIC code or the SCC.

In the data merge, where data were supplied for the same facility, MACT category, and pollutant from two or more data sources, only one data source was chosen. In choosing which records to keep, the hierarchy for the most part was as follows:

- State and local agency-supplied data;
- MACT standard analyses data; and
- TRI data.

Table III-3 indicates the relative percentages of the data sources for this inventory. As shown below, the overwhelming majority of the data is state data submitted data for 2002 (approximately 88 percent). Data from the TRI inventory accounts for approximately 12 percent of the point source benzene emissions.

The 2002 EGU data (noted as 767/CAMD, 767/CAMD1, and CAMD data sources) are a combination of data from CAMD (EPA, 2004a), DOE (DOE, 2003), and EPA-approved emission factors (ERG and Pechan, 2004).

Table III-3. Data Sources for 2002 Houston-Galveston Point Source Benzene Emissions

Data Source	Data Source Definition	Year	2002 Benzene Emissions (tpy)	% of Total
S	State Submitted	2002	673.91	87.19
M	EPA's ESD for MACT/residual risk source categories (only MACT code 0603 - Marine Vessel Loading Operations was identified with this data source).	1999	6.45	0.83
T	TRI	2002	91.82	11.88
767/CAMD	Record in both 2002 Form EIA-767 and 2002 Emission tracking system/continuous emissions monitoring (ETS/CEM); with sulfur dioxide (SO_2), oxides of nitrogen (NO_x), and heat input values from ETS/CEM; condensible PM, and primary PM_{10} and $PM_{2.5}$ are recalculated using CAMD heat input values.	2002	0.33	0.04
767/CAMD1	Combined cycle record (heat recovery steam generators + combustion turbine) in 2002 ETS/CEM used; the heat recovery steam generators record in 2002 Form EIA-767 eliminated. SO_2, NO_x, and heat input values from ETS/CEM; other emissions estimated.	2002	0.20	0.03
CAMD	Record only in 2002 ETS/CEM for SO_2, NO_x, and heat input values; other emissions estimated.	2002	0.17	0.02
Total		2002	772.88	100.00

ii. 2002 Database Augmentations

Economic growth activity factors are linked with the point source database using MACT codes and SCC codes. In the 2002 NEI point source file, a number of facility process records were missing, or had invalid, SCC codes. In addition, a number of records had composite MACT codes (MLTxxx) – indicating that EPA, in the preparation of the inventory, found that the facility may be affected by more than one MACT standard.

a. Assigning SCCs

In order to perform the 2010 and 2020 *with-CAAA* projections, any point source records with missing SCC codes in the 2002 NEI were augmented. These assignments were made using available SIC and MACT codes. Table III-4 below indicates the crosswalk assignments for the 26 emission records with missing SCCs.

Table III-4. SIC/MACT to SCC Crosswalk for 2002 Draft NEI Point Source File

Process MACT Code	MACT Source Category Name	SIC	SIC Description	Assigned SCC	SCC Description
		2992	Lubricating Oils and Greases	30610001	Lube Oil Refining
		4953	Refuse Systems	40290023	Fuel-fired Equipment-Flares
0560	Cyclic Crude and Intermediate Production	2865	Cyclic Organic Crudes and Intermediates	30101904	Phthalic Anhydride
0601	Gasoline Dist. (Stage I)	5171	Petroleum Bulk Stations and Terminals	40400406	Underground Tanks: RVP7 Gasoline
0602	Organic Liquids Dist. (non-gasoline)	2869	Chem. Mfg.-Adipic Acid	40721205	Carbitol: Standing loss
1461	Indus. Organic Chem. Manuf.	2819	Indus. Organic Chemicals	30101101	Hydrochloric Acid-Byproduct
1560		2869	Chem. Mfg.-Adipic Acid	30100101	Adipic Acid General
1641	Misc. Organic Chem. Manuf.	2841	Chemicals & Allied Products-Detergents	30100901	Spray Drying-Soaps and Detergents
1641	Misc. Organic Chem. Manuf.	2869	Indus. Organic Chemicals	30103402	Aniline/Ethanolamines

In addition, there were 28 emission records with inactive SCCs. These SCCs were mapped according to EPA direction in the NIF SCC reference table. Table III-5 provides this mapping.

Table III-5. Invalid SCC Mapping for 2002 Draft Point NEI

Original SCC	Map To SCC	SCC Description
28888802	28888801	Internal Combustion Engines : Fugitive Emissions : Other Not Classified : Specify in Comments
30688802	30688801	Industrial Processes : Petroleum Industry : Fugitive Emissions : Specify in Comments Field
30688803	30688801	Industrial Processes : Petroleum Industry : Fugitive Emissions : Specify in Comments Field
30688804	30688801	Industrial Processes : Petroleum Industry : Fugitive Emissions : Specify in Comments Field
30688805	30688801	Industrial Processes : Petroleum Industry : Fugitive Emissions : Specify in Comments Field
30699998	30699999	Industrial Processes : Petroleum Industry : Petroleum Products - Not Classified : Not Classified **
30800197	30800199	Industrial Processes : Rubber and Miscellaneous Plastics Products : Tire Manufacture : Other Not Classified
40388802	40388801	Petroleum and Solvent Evaporation : Petroleum Product Storage at Refineries : Fugitive Emissions : Specify in Comments Field
40388805	40388801	Petroleum and Solvent Evaporation : Petroleum Product Storage at Refineries : Fugitive Emissions : Specify in Comments Field

The above SCC assignments for missing or invalid values affected approximately 92 tpy of benzene emissions. Table III-6 identifies these emissions by source category.

Table III-6. Tier 2 Summary of Assigned SCC Emissions for the 2002 Draft NEI Point Source File

TIER 1	TIER 1 NAME	TIER 2	TIER 2 NAME	Associated Benzene Emissions (tpy)
02	FUEL COMB. INDUSTRIAL	03	Gas	0.31
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	34.83
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	1.53
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	2.85
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	0.05
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	1.23
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	51.03
			Total	**91.82**

b. MACT Code Assignments

To evaluate EPA's progress in reducing air toxic emissions through the MACT standards and to identify sources that may be modeled as part of residual risk assessments, operations within facilities that are subject to MACT standards are identified in the NEI by MACT codes. The tagging of data with MACT codes allows EPA to determine reductions attributable to the MACT program. The NEI associates MACT codes with major and area sources. MACT codes are assigned at the process level, or at the site level, in the point source file, e.g., the MACT code for municipal waste combustors is assigned at the site level, whereas the MACT code for petroleum refinery catalytic cracking is assigned at the process level. MACT codes are also assigned to source categories in the nonpoint source file. EPA requested that State and local agencies and tribes include MACT codes as part of their submittal of 2002 HAP emission inventory data. If they did not include MACT codes in their inventories, EPA assigned them, where appropriate.

Records with composite MACT codes were identified as composites of the following values as defined by the EPA and listed in Table III-7. These are situations where EPA believes that more than one MACT standard may affect the point source.

Table III-7. Multiple MACT Code Listing

Multiple MACT	MACT Code 1	MACT Code 1 Description	MACT Code 2	MACT Code 2 Description	MACT Code 3	MACT Code 3 Description	MACT Code 4	MACT Code 4 Description
MLT102	1405	Cyanide Chemicals Manufacturing	1641	Miscellaneous Organic Chemical Manufacturing				
MLT105	1409	Hydrogen Fluoride Production	1641	Miscellaneous Organic Chemical Manufacturing				
MLT107	1461	Industrial Inorganic Chemical Manufacturing	1641	Miscellaneous Organic Chemical Manufacturing				
MLT29	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	0503	Petroleum Refineries - Other Sources Not Distinctly Listed				
MLT33	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	0503	Petroleum Refineries - Other Sources Not Distinctly Listed	0602	Organic Liquids Distribution (Non-Gasoline)		
MLT34	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	0503	Petroleum Refineries - Other Sources Not Distinctly Listed	0602	Organic Liquids Distribution (Non-Gasoline)	1635	Ethylene Processes
MLT36	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	0503	Petroleum Refineries - Other Sources Not Distinctly Listed	1635	Ethylene Processes		
MLT47	0602	Organic Liquids Distribution (Non-Gasoline)	1635	Ethylene Processes				
MLT49	0602	Organic Liquids Distribution (Non-Gasoline)	1641	Miscellaneous Organic Chemical Manufacturing				
MLT77	0911	Pesticide Active Ingredient Production	1461	Industrial Inorganic Chemical Manufacturing				

The records with composite MACT codes were replaced by assigning the following values (based on EPA documentation of the composition of the multiple MACT code values) as shown in Table III-8. MACT Code 1 from Table III-7 was assigned as the primary MACT code in all cases.

Table III-8. Summary of Multiple MACT Code Assignments

Original Multiple MACT Code	Defaulted MACT Code	MACT Description	Associated Benzene Emissions (tpy)
MLT29	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	42.82
MLT33	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	18.56
MLT34	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	4.76
MLT36	0502	Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units	2.64
MLT47	0602	Organic Liquids Distribution (Non-Gasoline)	15.88
MLT49	0602	Organic Liquids Distribution (Non-Gasoline)	1.58
MLT77	0911	Pesticide Active Ingredient Production	0.04
MLT102	1405	Cyanide Chemicals Manufacturing	0.83
MLT105	1409	Hydrogen Fluoride Production	0.60
MLT107	1461	Industrial Inorganic Chemical Manufacturing	0.00
		Total	**87.69**

The majority of emissions with defaulted MACT codes were those affecting the following two source categories: (1) Petroleum Refineries - Catalytic Cracking, Catalytic Reforming, & Sulfur Plant Units; and (2) Organic Liquids Distribution (Non-Gasoline).

c. Control Device Information

In designing an approach for examining the effects of current and expected future control levels on benzene emissions, the information in the 2002 draft NEI point source file was reviewed. The database field that indicates whether a source has existing pollution controls provided the information summarized in Table III-9. (This field is not used for calculation; it is only a qualitative indicator.)

Table III-9. Summary of Control Status for 2002 Point Sources

Control Status	2002 Benzene Emissions (tpy)
CONTROLLED	219.22
UNCONTROLLED	461.84
UNKNOWN	91.82
Total	**772.88**

However, when examining the presence or absence of benzene control efficiencies in the NEI (total capture control efficiency/primary percent control efficiency), it was found that only two facilities had identified non-zero or non-null control efficiencies. In addition, the sum of emissions for these two facilities was less than one ton of benzene emissions. The remaining control records identified control devices, but not the associated control efficiencies. This finding prompted the decision to develop 2010 and 2020 forecasts using only MACT-indicated control efficiencies from a 2002 baseline for application in the emission projections.

b. Base Year Source for without-CAAA Projections

i. 1990 NEI for HAPs

The 1990 NEI for HAPs point source file (EPA, 2005a) was used as the base year for the *without-CAAA* projections. The original baseline 1990 NTI was a county-level inventory for all source categories. The newly released 1990 NEI for HAPs was created by converting the county-level emission estimates to facility-specific estimates for as many sources as possible. Locational data and stack parameters were added. Additional estimates were developed for missing MACT source categories and HAPs so that the baseline inventory is more comparable to the 1999 NEI and the 2002 NEI (EPA, 2005b).

This section of the report presents a summary of the 1990 Point NEI based on the EPA Tier level grouping of source categories. Table III-10 displays the annual emissions of 585 facilities in the Brazoria-Galveston-Harris County study area.

Table III-10. Tier 2 Summary of 1990 NEI Point Emission Inventory

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	1990 Benzene Emissions (tpy)	% of Total
01	FUEL COMB. ELEC. UTIL.	02	Oil	0.00	0.00
01	FUEL COMB. ELEC. UTIL.	03	Gas	0.08	0.00
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.00	0.00
02	FUEL COMB. INDUSTRIAL	03	Gas	27.15	1.04
02	FUEL COMB. INDUSTRIAL	04	Other	0.03	0.00
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	7.27	0.28
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.01	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	450.91	17.32
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	57.14	2.20
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	91.99	3.53
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	296.27	11.38
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	44.70	1.72
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	826.07	31.74
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	90.34	3.47
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	16.00	0.61
07	OTHER INDUSTRIAL PROCESSES	08	Transportation Equipment	8.12	0.31
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	61.08	2.35
08	SOLVENT UTILIZATION	04	Surface Coating	1.63	0.06
08	SOLVENT UTILIZATION	05	Other Industrial	0.06	0.00
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	51.11	1.96
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	127.96	4.92
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	191.71	7.36
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	199.52	7.67
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	6.37	0.24
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	2.00	0.08
10	WASTE DISPOSAL & RECYCLING	03	POTW	7.68	0.30
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	33.51	1.29
10	WASTE DISPOSAL & RECYCLING	05	TSDF	4.28	0.16
			Total	**2602.99**	**100.00**

In 1990, five categories contribute approximately 75 percent of the emissions:

- Chemical & Allied Product Mfg - Organic Chemical Mfg (17.3 percent);
- Petroleum & Related Industries - Petroleum Refineries & Related Industries (31.7 percent);
- Storage & Transport - Organic Chemical Storage (7.7 percent);
- Storage & Transport - Petroleum & Petroleum Product Storage (7.4 percent); and
- Chemical & Allied Product Mfg - Other Chemical Mfg (11.4 percent).

These priority source categories are the same as the top five categories in the 2002 Point NEI, with some differences in the percentage breakdowns.

Key data sources for the 1990 Point NEI are shown in Table III-11.

Definitions for these data sources were received in conversation with the inventory developer. Over half of the emissions are identified as *Local*, whereas approximately one-third are identified as *ESD-alloc*. The remainder of the emissions is identified as *Transfer* or *Calc-alloc*.

The EPA established a hierarchy of preferred data sources in order to prepare the 1990 NEI for HAPs, listed below in order of preference:

- MACT data from EPA’s Emission Standards Division (ESD);
- Data developed by state and local air agencies;
- Data from inventories developed by EPA’s Emission Inventory Group to support requirements of Sections 112(c)(6) and 112(k); and
- Emissions reported in the Toxic Release Inventory (TRI), and emissions that the Emission Inventory Group generated using emission factors and activity factors.

Emissions data for Harris County, Texas were among the state/local agency data that were available at the time that EPA was compiling the initial baseline.

Table III-11. Data Sources for 1990 Point NEI Benzene Emission Estimates Houston-Galveston Study Area

Data Source	Data Source Definition	1990 Benzene Emissions (tpy)	% of Total
CALC-alloc	Emissions were calculated and then allocated based on an activity factor, usually capacity	92.61	3.56
ESD-alloc	Emissions were calculated based on ESD data and then allocated based on an activity factor, usually capacity.	852.71	32.76
LOCAL	State or local data	1456.82	55.97
TRANSFER	Data originated from TRI database, ESD database, or 1996 NEI for point sources	200.86	7.72
Total		2603	100.00

The source categories in the baseline 1990 NEI (1,012 reported categories, including mobile sources) were evaluated to determine the feasibility of converting county-level estimates to a point source inventory. In order to do this, individual facilities with associated emissions data had to be defined. TRI source categories contained point source level information and were included in the point sources inventory. There were some non-TRI based source categories that held information at the facility-level; however, these were researched before they were included in the point source inventory. If a source category did not have point source locational elements, they were suited for inclusion in the nonpoint inventory. Mobile source county-level emission estimates were directly transferred for all onroad and nonroad sources. Some source categories contained some point source locational elements, and were included in both point and nonpoint inventories. Some source categories were added for completeness (when compared to later year inventories). An example of this category would be gold mining.

Since the 1990 NEI for HAPs was a compilation of data from multiple sources, an overlap analysis was performed. The point source inventory was reviewed for unique facilities and pollutants. An analysis was also performed to review for potential overlaps between point source categories and nonpoint source categories.

ii. 1990 Database Augmentations

Assigning SCCs

A number of facility process records did not include valid SCCs in the 1990 point source benzene inventory. As described below, Pechan assigned a valid SCC to each of these records in order to have the necessary information for applying growth and control factors and preparing emission summaries by source category.

Where an emission record was missing an SCC, Pechan generally assigned an SCC based on the facility name, emission process description, and SIC code information reported for that record. In cases where more than one relevant SCC existed in the inventory, Pechan selected the SCC with the highest benzene emissions. For example, there are two emission process records for the facility entry “PHILLIPS 66 CO. SAN BERNARD TERMINAL,” each of which is associated with a process description of “Petroleum Bulk Stations and Terminals” and an SIC code of 5171 (Petroleum Bulk Stations and Terminals). Based on this information, Pechan reviewed the 1990 point source emissions for SCCs associated with Bulk Terminals, and assigned these records to SCC 40400170 − Petroleum Liquids Storage (non-Refinery); Bulk Terminals; Specify Liquid: Standing Loss-Internal Floating Roof with Secondary Seal because this SCC accounted for the highest Bulk Terminal benzene emissions (16.7 tpy) in the 1990 point source inventory. For some emission records, SCCs that matched to facility name, emission process description, and SIC code did not appear in the 1990 inventory. For these records, Pechan generally selected the SCC that reflected the most general process.

There were 23 emission records with an invalid SCC in the point source inventory. Pechan used two approaches for assigning a valid SCC for these records. Whenever possible, Pechan assigned the record to the new SCC that EPA identifies as the replacement for the invalid SCC in their latest official SCC list. This approach was used for three emission records. The remaining 20 emission records were reported in the 1990 inventory as having 6-digit SCCs (valid point SCCs require 8 digits). For these records, Pechan assigned an 8-digit SCC that both fell within the reported 6-digit SCC and for which a benzene emission factor was available from EPA’s

WebFIRE (EPA, 2006). When more than one SCC fit these criteria, Pechan selected the most general SCC.

Houston-Galveston 1990 point sources with approximately 197 tpy of benzene emissions had SCCs added or modified. Table III-12 identifies the SCC assignments for the 68 point source emission records with missing or invalid SCCs in the 1990 NEI point source file. Table III-13 summarizes these emissions based on their Tier 2 source categories.

MACT Code Assignments
No MACT code assignment changes were made for the 1990 Point NEI.

Missing Control Information
No control information was provided with the 1990 data; all records indicated an uncontrolled status.

B. 1990 VERSUS 2002 TOP EMITTING FACILITY COMPARISON

As part of the case study analysis, Pechan evaluated the facility and unit-specific data for the top emitting facilities from 1990 NEI and 2002 NEI. The purpose of this comparison was to try to identify the major influences on benzene emission changes during this recent historical period. Note that facilities were matched primarily based on NTI Site ID (as it was not possible to match on the state provided facility identifier). In some cases the NTI Site ID was not available. Table III-14 summarizes some of the emissions and identifier information for these top facilities.

For the top 10 benzene-emitting point sources in the 1990 NEI for HAPs, it was found that six facilities' emissions and supporting data were provided by local sources (i.e., a point source data file for Harris County that was provided by the TCEQ). Point source records for the other four facilities in the 1990 top 10 were based on the Emission Standards Division allocation method. The six local data facilities are each represented by a single record in the 1990 database. Therefore, for these six facilities, there was not enough information about 1990 operations to determine the major influences on benzene emissions between 1990 and 2002. Another issue with the 1990 records for these facilities (which are major petrochemical plants) is that all of the benzene emissions will be modeled as if they are all emitted from a single point within the facilities. For any nearby monitors, this treatment will not be as accurate as having identified the correct emission locations and stack parameters for the major benzene sources within these facilities.

For the four 1990 top benzene emitting point sources whose 1990 emissions were estimated using the Emission Standards Division allocation method, it was found that 1990 emissions were almost always 8.33 times 2002 emissions. These are all major chemical manufacturing facilities where it appears that EPA has back calculated 1990 benzene emissions using a simple factor to account for the estimated effects of the MACT standards during this period. Unfortunately, the effects of the MACT standard (which is equivalent to an 88 percent reduction in benzene) has been applied equally across all benzene sources in each facility, so the differential effects of the MACT requirements on different sources within a plant has not been captured.

Table III-12. Crosswalk for Emission Records with Missing/Invalid SCCs in the 1990 Point Source Inventory

Facility Name	Emission Unit ID	Process ID	Emission Process Description	SIC Unit Level	SIC Description	Assigned SCC	SCC Description
PHILLIPS 66 CO. SAN BERNARD TERMINAL	1	1	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
PHILLIPS 66 CO. SAN BERNARD TERMINAL	1	2	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
PHILLIPS 66 CO. FREEPORT I TERMINAL	1	1	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
PHILLIPS 66 CO. FREEPORT I TERMINAL	1	2	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
PHILLIPS 66 CO. JONES CREEK TERMINAL	1	1	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
PHILLIPS 66 CO. JONES CREEK TERMINAL	1	2	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
OXYCHEM PETROCHEMICALS INC.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
OXYCHEM PETROCHEMICALS INC.	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
PHILLIPS 66 CO. FREEPORT TERMINAL II	1	1	Petroleum Bulk Stations and Terminals	5171	Petroleum Bulk Stations & Terminals	40400170	Bulk Terminals : Specify Liquid: Standing Loss - Int. Floating Roof w/ Secondary Seal
UNION CARBIDE CORP. MARINE TERMINAL	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
UNION CARBIDE CORP. MARINE TERMINAL	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
ISP TECHNOLOGIES INC.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
BLENTECH CORP.	1	1	Industrial Inorganic Chemicals, NEC	2819	Industrial Inorganic Chemicals	30187598	Inorganic Chemical Storage (Floating Roof Tank) : Specify Liquid: Withdrawal Loss
BLENTECH CORP.	1	2	Industrial Inorganic Chemicals, NEC	2819	Industrial Inorganic Chemicals	30187598	Inorganic Chemical Storage (Floating Roof Tank) : Specify Liquid: Withdrawal Loss
PILOT IND. OF TEXAS INC.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
PILOT IND. OF TEXAS INC.	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
TEXAS ALKYLS INC.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
TEXAS ALKYLS INC.	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
LUBRIPAC	1	1	Lubricating Oils and Greases	2992	Lubricating Oils And Greases	30630007	Re-refining of Lube Oils and Greases : Finished Product Storage Tank
LUBRIPAC	1	2	Lubricating Oils and Greases	2992	Lubricating Oils And Greases	30630007	Re-refining of Lube Oils and Greases : Finished Product Storage Tank

Table III-12 (continued)

Facility Name	Emission Unit ID	Process ID	Emission Process Description	SIC Unit Level	SIC Description	Assigned SCC	SCC Description
KOPPERS IND. INC. HOUSTON TAR PLANT	1	1	Cyclic Organic Crudes and Intermediates, and Organic Dyes and Pigments	2865	Industrial Organic Chemicals Cyclic Crudes And Intermediate	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
KOPPERS IND. INC. HOUSTON TAR PLANT	1	2	Cyclic Organic Crudes and Intermediates, and Organic Dyes and Pigments	2865	Industrial Organic Chemicals Cyclic Crudes And Intermediate	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
AMERICAN TEXMARK INC. DBA TEXMARK	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
AMERICAN TEXMARK INC. DBA TEXMARK	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
AMOCO CHEMICAL CO. CHOCOLATE BAYOU PLANT	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
AMOCO CHEMICAL CO. CHOCOLATE BAYOU PLANT	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
BASF CORP.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
BASF CORP.	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
DOW CHEMICAL CO. TEXAS OPERATIONS	1	1	Alkalies and Chlorine	2812	Alkalies And Chlorine	30100899	Chloro-alkali Production : Other Not Classified
DOW CHEMICAL CO. TEXAS OPERATIONS	1	2	Alkalies and Chlorine	2812	Alkalies And Chlorine	30100899	Chloro-alkali Production : Other Not Classified
MONSANTO CO.	1	1	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
MONSANTO CO.	1	2	Industrial Organic Chemicals, NEC	2869	Industrial Organic Chemicals,nec	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
STERLING CHEMICALS INC.	1	1	Cyclic Organic Crudes and Intermediates, and Organic Dyes and Pigments	2865	Industrial Organic Chemicals Cyclic Crudes And Intermediate	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
STERLING CHEMICALS INC.	1	2	Cyclic Organic Crudes and Intermediates, and Organic Dyes and Pigments	2865	Industrial Organic Chemicals Cyclic Crudes And Intermediate	40799997	Organic Chemical Storage : Miscellaneous : Specify in Comments
CHAMPION INTERNATIONAL	3A	3	Pulp and Paper Production (Non-Combustion) MACT I			30700199	Sulfate (Kraft) Pulping, Other Not Classified

Table III-12 (continued)

Facility Name	Emission Unit ID	Process ID	Emission Process Description	SIC Unit Level	SIC Description	Assigned SCC	SCC Description
CHAMPION INTERNATIONAL	5A	5	Pulp and Paper: Combustion			30700199	Sulfate (Kraft) Pulping, Other Not Classified
SIMPSON PAPER CO.	3A	3	Pulp and Paper Production (Non-Combustion) MACT I			30700199	Sulfate (Kraft) Pulping, Other Not Classified
SIMPSON PAPER CO.	5A	5	Pulp and Paper: Combustion			30700104	Sulfate (Kraft) Pulping, Recovery Furnace/Direct Contact Evaporator
SOUTHWESTERN BARGE FLEET SERVICE INC.	1	1	Ship Building and Repairing	3731	Ship Building And Repairing	31401501	Industrial Processes : Transportation Equipment : Boat Manufacturing : General
SOUTHWESTERN BARGE FLEET SERVICE INC.	1	2	Ship Building and Repairing	3731	Ship Building And Repairing	31401501	Industrial Processes : Transportation Equipment : Boat Manufacturing : General
CLEAN HARBORS DEER PARK, L.P.	1	1		4953	Sanitary Services Refuse Systems	50300801	Solid Waste Disposal - Industrial : Treatment, Storage, Disposal/TSDF : Surface Impoundment: Fugitive Emissions
CLEAN HARBORS DEER PARK, L.P.	1	2		4953	Sanitary Services Refuse Systems	50300801	Solid Waste Disposal - Industrial : Treatment, Storage, Disposal/TSDF : Surface Impoundment: Fugitive Emissions
E.I. DU PONT DE NEMOURS & COMPANY/HARRIS/HG0218K	1	1	Continuous	3081	Miscellaneous Plastics Products, Nec Unsupported Plastics Film And Sheet	30101861	Plastics Production, Purification System (Polyethylene)
QUANTUM CHEMICAL CORPORATION/HARRIS/HG0770G	1	1	Continuous	2821	Synthetics Plastics Materials And Resins	30101861	Plastics Production, Purification System (Polyethylene)
SOLVAY POLYMERS, INC./HARRIS/HG0665E	1	1	Continuous	2821	Plastics Materials And Resins	30101861	Plastics Production, Purification System (Polyethylene)
P ROBINSON	FO6	ST	Heavy Oil - Steam			10100401	Electric Generation : Residual Oil : Grade 6 Oil: Normal Firing
GREEN BAYOU	FO6	ST	Heavy Oil - Steam			10100401	Electric Generation : Residual Oil : Grade 6 Oil: Normal Firing
SAM BERTRON	FO6	ST	Heavy Oil - Steam			10100401	Electric Generation : Residual Oil : Grade 6 Oil: Normal Firing
T H WHARTON	FO6	ST	Heavy Oil - Steam			10100401	Electric Generation : Residual Oil : Grade 6 Oil: Normal Firing
P ROBINSON	NG1	ST	Natural Gas - Steam	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
GREEN BAYOU	NG2	ST	Natural Gas - Steam	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
DEEPWATER	NG	ST	Natural Gas - Steam			10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
T H WHARTON	NG3	ST	Natural Gas - Steam	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
SAM BERTRON	NG1	ST	Natural Gas - Steam	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential

Table III-12 (continued)

Facility Name	Emission Unit ID	Process ID	Emission Process Description	SIC Unit Level	SIC Description	Assigned SCC	SCC Description
WEBSTER	NG2	ST	Natural Gas - Steam	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
T H WHARTON	NG3	CT	Natural Gas - Combined Cycle (gas)	4911	Electric Services	10100601	Electric Generation : Natural Gas : Boilers > 100 Million Btu/hr except Tangential
WEBSTER	FO2	IC	Light Oil - Internal Combustion			20100102	Electric Generation : Distillate Oil (Diesel) : Reciprocating
T H WHARTON	FO2	GT	Light Oil - Gas Turbine			20100101	Electric Generation : Distillate Oil (Diesel) : Turbine
GREEN BAYOU	FO2	GT	Light Oil - Gas Turbine			20100101	Electric Generation : Distillate Oil (Diesel) : Turbine
P ROBINSON	NG1	GT	Natural Gas - Gas Turbine	4911	Electric Services	20100201	Electric Generation : Natural Gas : Turbine
GREEN BAYOU	NG2	GT	Natural Gas - Gas Turbine	4911	Electric Services	20100201	Electric Generation : Natural Gas : Turbine
WEBSTER	NG2	GT	Natural Gas - Gas Turbine	4911	Electric Services	20100201	Electric Generation : Natural Gas : Turbine
SAM BERTRON	NG1	GT	Natural Gas - Gas Turbine	4911	Electric Services	20100201	Electric Generation : Natural Gas : Turbine
T H WHARTON	NG3	GT	Natural Gas - Gas Turbine	4911	Electric Services	20100201	Electric Generation : Natural Gas : Turbine
H O CLARKE	NG	GT	Natural Gas - Gas Turbine			20100201	Electric Generation : Natural Gas : Turbine
Houston Pipeline Co.	53	1	Electric, Gas, And Sanitary Services, Gas Production and Distribution			28888801	Fugitive Emissions : Other Not Classified : Specify in Comments
Houston Pipeline Co.	54	1	Electric, Gas, And Sanitary Services, Gas Production and Distribution			28888801	Fugitive Emissions : Other Not Classified : Specify in Comments
Houston Pipeline Co.	56	1	Electric, Gas, And Sanitary Services, Gas Production and Distribution			28888801	Fugitive Emissions : Other Not Classified : Specify in Comments

NOTE: 1990 point source emission records that had invalid SCCs are noted in shading

.

Table III-13. Tier 2 Summary of Assigned SCC Emissions for the 1990 Point NEI

TIER 1	TIER 1 NAME	TIER 2	TIER 2 NAME	Associated Benzene Emissions (tpy)
01	FUEL COMB. ELEC. UTIL.	02	Oil	0.00
01	FUEL COMB. ELEC. UTIL.	03	Gas	0.08
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.00
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.01
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	0.14
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	91.82
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	16.78
07	OTHER INDUSTRIAL PROCESSES	08	Transportation Equipment	8.13
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	11.41
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	0.13
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	68.21
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	0.01
10	WASTE DISPOSAL & RECYCLING	05	TSDF	0.38
			Total	**197.09**

Table III-14. Facility Comparison - 1990 NEI to draft 2002 NEI

FIPS State/County	1990 ID	1990 NTI ID	1990 Facility	1990	2002	% 2002 to 1990	Match Type	2002 ID	2002 NTI	2002 Facility
48201	ESD1501-245	NEITXT$11693	CHANNELVIEW COMPLEX	418.02	50.16	12.00	NTI Site ID and Facility Name	75	NEITXT$11693	CHANNELVIEW COMPLEX
48167	853-122	NEI2TXT17872	Amoco Oil Co.	188.51	0.96	0.51	NTI Site ID	ES0603-127 ES0603-128	NEI2TXT17872	Amoco Oil Co., Dock No. 32. Amoco Oil Co., Docks Nos. 37 and 38.
48201	853-128		Exxon Co. USA	172.40	28.70	16.65	Facility Name	27	NEI7781	EXXONMOBIL REF & SUPPLY
48039	ESD1501-219	NEI11295	PLANT B	148.09	32.97	22.27	NTI Site ID and Facility Name	41	NEI11295	PLANT B
48201	853-130		Lyondell-Citgo Refining Co.	115.37	43.62	37.81	Facility Name	40	NEI11232	LYONDELL-CITGO REFINING
48039	ESD1501-211	NEITXT$11677	CHOCOLATE BAYOU CHEMICAL	99.31	39.60	39.87	NTI Site ID and Facility Name	48	NEITXT$11677	CHOCOLATE BAYOU CHEMICAL
48201	853-127		Deer Park Refining Limited Partnership	93.99	86.34	91.85	Probable Match on Facility Name	39	NEI11119	DEER PARK PLANT
48167	ESD1501-233	NEI6503	TEXAS CITY PLANT	85.88	11.13	12.96	NTI Site ID and Facility Name	10	NEI6503	TEXAS CITY PLANT
48039	853-119		Phillips 66 Co.	76.19	17.83	23.40	Probable Match on Location and Facility Name	10	NEI6519	SWEENY REFINERY PETROCHEM
48201	HG0674D-020	NEI2TX201144	Champion International Corporation	72.60		0.00	unable to match			

C. ANALYSIS OF ALTERNATIVE BENZENE DATA EMISSION SOURCES

While the 2002 draft Point NEI was selected as the base year inventory for the *with-CAAA* projection calculation and analysis and the 1990 NEI Point Inventory for HAPs was selected for the base year inventory for the *without-CAAA* projection calculations for point sources, other inventories were reviewed for benzene emissions in the Brazoria-Galveston-Harris counties as an indicator of the potential levels of uncertainties that may exist in the data sets that were selected.

The following inventories were also reviewed in order to provide a comparison to the 2002 draft Point NEI and provide an estimate of any uncertainty in the emission totals that may exist:

- 1996 Point National Toxics Inventory (NTI);
- 1999 Final Point NEI;
- 2002 TRI; and
- 2000 Texas AQS Inventory.

Tier 2 emission summaries of these data sets (where possible) and facility counts are provided in the following sections. In summary, the 1999 final point NEI and the 2002 TRI compare favorably with the 2002 draft Point NEI. The 1999 point NEI benzene emissions differed by approximately 3 percent; the 2002 TRI differed by 16 percent.

The 1996 inventory differed by almost a factor 2 – however, given the time difference (6 years), potential differences in reporting criteria at the State level and activity levels, there are a number of factors which may contribute to this difference.

The 2000 Texas AQS inventory differed by almost a factor of 3 – however, given that the original data were measured as daily values for peak ozone season day, extrapolating to an annual value by multiplying by 365, is certain to overestimate annual values.

The number of facilities was relatively consistent between the 1999 and 2002 inventories; however, there were wide gaps in the counts for 2000 Texas AQS, 1996 Point NEI, and the TRI inventory. Possible reasons for this variance are provided below.

In summary, the majority contributors to the emissions remained consistent across all the inventories. These majority contributors are as follows:

- Chemical & Allied Product Mfg - Organic Chemical Mfg;
- Petroleum & Related Industries - Petroleum Refineries & Related Industries;
- Storage & Transport - Organic Chemical Storage;
- Storage & Transport - Petroleum & Petroleum Product Storage; and
- Chemical & Allied Product Mfg - Other Chemical Mfg.

There are some variations among data sets as noted below.

1. 1996 NTI

The 1996 point source NTI is of interest because it was used in a previous EPA analysis of benzene emissions, concentrations, and associated risk for the Houston-Galveston area. The 1996 NTI emissions summary shown in Table III-15 represent the annual emissions of 2,224 facilities in the Brazoria-Galveston-Harris county area.

In comparing this inventory to the 2002 Point NEI, there are two points to note. The number of facilities is considerably different (by an order of magnitude, compared to the 199 facilities for 2002) and the 1996 benzene emissions are nearly a factor of two higher than in the 2002 NEI. In addition, approximately 40 percent of these emissions are unclassified according to the current Tier Level – SCC mapping. These emissions are unclassifiable because the emissions do not have an associated SCC.

The top benzene emitting categories are also somewhat different from those in the 2002 NEI. In the 1996 NTI, the following categories account for approximately 50 percent of the emissions (noting that 40 percent of the total emissions are unclassifiable, therefore, together these emissions represent more than 90 percent of the emissions):

- Chemical & Allied Product Mfg - Organic Chemical Mfg (14.86 percent);
- Petroleum & Related Industries - Petroleum Refineries & Related Industries (9.07 percent);
- Storage & Transport - Petroleum & Petroleum Product Storage (8.68 percent);
- Chemical & Allied Product Mfg - Other Chemical Mfg (3.32 percent);
- Storage & Transport - Organic Chemical Storage (3.68 percent);
- Waste Disposal & Recycling – POTW (4.51 percent); and
- Storage & Transport - Petroleum & Petroleum Product Transport (2.90 percent).

The breakdown in top categories as compared to the 2002 draft Point NEI is similar with two exceptions:

- Waste Disposal & Recycling – POTW (4.51 percent); and
- Storage & Transport - Petroleum & Petroleum Product Transport (2.90 percent).

Of these two, the difference is significant for Waste Disposal & Recycling – POTW. This category accounts for a significant percentage in the 2002 draft Point NEI inventory for Brazoria-Galveston-Harris counties for benzene.

Table III-15. Tier 2 Summary of 1996 NTI Point Source Houston-Galveston Study Area

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	1996 Benzene Emissions (tpy)	% of Total
	Unclassified			562.66	42.91
01	FUEL COMB. ELEC. UTIL.	03	Gas	0.08	0.01
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.17	0.01
02	FUEL COMB. INDUSTRIAL	02	Oil	0	0.00
02	FUEL COMB. INDUSTRIAL	03	Gas	51.57	3.93
02	FUEL COMB. INDUSTRIAL	04	Other	0	0.00
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	7.26	0.55
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.01	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	194.83	14.86
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	0.47	0.04
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	0.3	0.02
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	43.58	3.32
05	METALS PROCESSING	01	Non-Ferrous Metals Processing	0.01	0.00
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	15.96	1.22
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	118.97	9.07
06	PETROLEUM & RELATED INDUSTRIES	03	Asphalt Manufacturing	0.38	0.03
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	7.56	0.58
07	OTHER INDUSTRIAL PROCESSES	04	Rubber & Miscellaneous Plastic Products	0	0.00
07	OTHER INDUSTRIAL PROCESSES	05	Mineral Products	0	0.00
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	2.98	0.23
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	0.81	0.06
08	SOLVENT UTILIZATION	03	Dry Cleaning	0.11	0.01
08	SOLVENT UTILIZATION	04	Surface Coating	0.65	0.05
08	SOLVENT UTILIZATION	05	Other Industrial	0.19	0.01
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	20.02	1.53
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	113.83	8.68
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	38.02	2.90
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	48.26	3.68
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	8.41	0.64
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	0.99	0.08
09	STORAGE & TRANSPORT	11	Bulk Materials Storage	0	0.00
10	WASTE DISPOSAL & RECYCLING	03	POTW	59.2	4.51
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	10.03	0.76
10	WASTE DISPOSAL & RECYCLING	05	TSDF	3.9	0.30
			Total	**1311.21**	**100.00**

Possible reasons for the differences between the 1996 NTI inventory and the 2002 NEI could include the following:

- The differences in facility counts and emission totals could result from data merged from another inventory (such as the TRI inventory set) into the 1996 inventory set without resolving potential double counting. The TRI inventory, for example, does not include SCC codes. When emissions without identifiable SCC codes are removed from the inventory – the total begins to approximate the 2002 Point NEI value. In addition – it is to be noted that the facility counts very closely approximate the emission process counts (2,224 to 2,324). It is possible that the facility identification by site identifier was oriented more to identification by process than the physical facility. Essentially, it seems that instead of identifying a physical facility by site identifier with multiple processes, a facility was given multiple site identifiers (possibly representing multiple emission units or processes).

- The differences in the distribution among the top categories could be the result of the including facilities as point sources in the 1996 NTI inventory that were aggregated and reported as nonpoint sources in the 2002 NEI.

2. 1999 NEI

For this analysis, the benzene emissions for Brazoria, Galveston, and Harris counties were extracted from the 1999 point NEI summary files (posted September 2003) at (http://www.epa.gov/ttn/chief/net/1999inventory.html). Table III-16 summarizes the annual emissions of 180 reporting facilities in the Brazoria-Galveston-Harris county area.

The primary contributors to the emissions are as follows (contributing approximately 90 percent of the emissions).

- Chemical & Allied Product Mfg - Organic Chemical Mfg (22.35 percent);
- Storage & Transport - Petroleum & Petroleum Product Storage (17.38 percent);
- Petroleum & Related Industries - Petroleum Refineries & Related Industries (17.26 percent);
- Chemical & Allied Product Mfg - Other Chemical Mfg (9.65 percent);
- Fuel Comb. Industrial - Gas (6.78 percent);
- Other Industrial Processes - Miscellaneous Industrial Processes (6.41 percent);
- Storage & Transport - Organic Chemical Storage (6.01 percent); and
- Storage & Transport - Petroleum & Petroleum Product Transport (4.40 percent).

Table III-16. Tier 2 Summary of 1999 NEI Point Source File Houston-Galveston Study Area

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	1999 Benzene Emissions (tpy)	% of Total
	Unclassified			0	0.00
01	FUEL COMB. ELEC. UTIL.	03	Gas	0	0.00
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.17	0.02
02	FUEL COMB. INDUSTRIAL	02	Oil	0.09	0.01
02	FUEL COMB. INDUSTRIAL	03	Gas	50.89	6.78
02	FUEL COMB. INDUSTRIAL	04	Other	4.44	0.59
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	5.55	0.74
03	FUEL COMB. OTHER	03	Commercial/Institutional Gas	0.26	0.03
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	167.69	22.35
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	0.15	0.02
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	0.02	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	04	Agricultural Chemical Mfg	0	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	72.41	9.65
05	METALS PROCESSING	02	Ferrous Metals Processing	2.21	0.29
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	13.26	1.77
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	129.49	17.26
06	PETROLEUM & RELATED INDUSTRIES	03	Asphalt Manufacturing	0.09	0.01
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	2.74	0.37
07	OTHER INDUSTRIAL PROCESSES	04	Rubber & Miscellaneous Plastic Products	0.1	0.01
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	2.04	0.27
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	48.11	6.41
08	SOLVENT UTILIZATION	03	Dry Cleaning	0.01	0.00
08	SOLVENT UTILIZATION	04	Surface Coating	1.29	0.17
08	SOLVENT UTILIZATION	05	Other Industrial	0.47	0.06
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	16.12	2.15
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	130.37	17.38
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	33.01	4.40
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	45.07	6.01
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	4.98	0.66
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	3.39	0.45
10	WASTE DISPOSAL & RECYCLING	03	POTW	7.84	1.05
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	7.75	1.03
10	WASTE DISPOSAL & RECYCLING	06	Landfills	0.01	0.00
10	WASTE DISPOSAL & RECYCLING	07	Other	0.21	0.03
			Total	**750.23**	**100.00**

The primary contributors to the emissions are similar to those in the other inventories. It is important to note that Industrial Combustion – Gas and Other Industrial Processes do contribute a relatively significant amount to this inventory whereas in the 2002 Point NEI they do not. This could occur for a number of reasons; the most likely include the following:

- The differences could be due to more specific SCC identification in the 2002 NEI (classifying processes under Chemical and Allied Process Manufacturing as opposed to Other Industrial Processes).

- Inclusion of emissions in the nonpoint emission inventory in the 2002 NEI where they may have previously been reporting as point sources.

Total benzene emissions are similar when compared with the 2002 NEI (750 tons in 1999 vs. 773 tons in 2002), and the total facility count is also similar (199 facilities vs. 180 facilities). Differences in facility counts could be due to facilities closing, potential removal of duplicate facilities from the 1999 NEI by either the submitting agency or EPA, or potentially low-reporting thresholds for some of the facilities.

3. 2002 TRI

The TRI is a publicly available EPA database that contains information on toxic chemical releases and other waste management activities reported annually by certain covered industry groups as well as federal facilities. This inventory was established under the Emergency Planning and Community Right-to-Know Act of 1986 and expanded by the Pollution Prevention Act of 1990. The TRI contains both HAP and ammonia emissions data. For this analysis, the benzene emissions for Brazoria, Galveston and Harris counties were extracted from the TRI inventory from the site http://www.epa.gov/tri/tridata/state_data_files.htm#TX.

The emissions in Table III-17 represent the annual emissions of 62 reporting facilities in the Brazoria-Galveston-Harris County area.

The TRI inventory does not contain SCC codes; therefore, a direct tier comparison is not available. However, using the SIC codes, it was found that approximately 85 percent of the TRI benzene emissions are from the following two industries:

- Chemicals and Allied Products - Industrial Organic Chemicals; and
- Petroleum Refining and Related Industries.

This is a similar percentage contribution from these two industries as was found in other inventories.

The number of independent facilities reporting to the TRI inventory is considerably less than the draft 2002 Point NEI (62 facilities vs. 180 facilities). The emissions however are similar (650 tons vs. 773 tons), indicating that either the identification of a unique facility may differ in TRI data set compared with the draft 2002 Point NEI, or that the TRI reporting requirements capture most of the facilities that are important benzene emitters in the area.

Table III-17. SIC Summary of 2002 TRI Emission Inventory Houston-Galveston Study Area

SIC	Description	2002 Benzene Emissions (tpy)	% Total
	Unclassified	0.01	0.00
2812	Chemicals And Allied Products Industrial Inorganic Chemicals Alkalis And Chlorine	33.54	5.16
2813	Chemicals And Allied Products Industrial Inorganic Chemicals Industrial Gases	1.48	0.23
2819	Chemicals And Allied Products Industrial Inorganic Chemicals Industrial Inorganic Chemicals	1.53	0.23
2841	Chemicals And Allied Products Soap, Cleaners, And Toilet Goods Soap And Other Detergents	2.85	0.44
2865	Chemicals And Allied Products Industrial Organic Chemicals Cyclic Crudes And Intermediate	47.77	7.35
2869	Chemicals And Allied Products Industrial Organic Chemicals Industrial Organic Chemicals, nec	322.85	49.64
2879	Chemicals And Allied Products Agricultural Chemicals Agricultural Chemicals, nec	0.00	0.00
2899	Chemicals And Allied Products Miscellaneous Chemical Products Chemical Preparations, Nec	1.58	0.24
2911	Petroleum Refining And Related Industries Petroleum Refining Petroleum Refining	233.23	35.86
2992	Petroleum Refining And Related Industries Misc Petroleum And Coal Products Lubricating Oils And Greases	0.05	0.01
3731	Transportation Equipment Ship And Boat Building And Repairing Ship Building And Repairing	1.80	0.28
4953	Electric, Gas And Sanitary Services Sanitary Services Refuse Systems	0.30	0.05
5171	Wholesale Trade-Nondurable Goods Petroleum And Petroleum Products Petroleum Bulk Stations & Terminals	3.35	0.52
	Total	**650.33**	**100.00**

4. 2000 Texas AQS Inventory

The following data set is from the Texas AQS inventory for year 2000. It is important to note that the original data were developed from daily emissions representing peak ozone season days. These data were developed to represent emissions during a very specific period – August 18, 2000 to September 7, 2000.

The emissions in Table III-18 represent the annual emissions of 320 reporting facilities in the Brazoria-Galveston-Harris county area.

Approximately 84 percent of the emissions are from the following categories:

- Storage & Transport - Petroleum & Petroleum Product Storage (28.35 percent);
- Chemical & Allied Product Mfg - Organic Chemical Mfg (19.31 percent);
- Chemical & Allied Product Mfg - Other Chemical Mfg (10.43 percent);
- Fuel Comb. Industrial - Gas (10.67 percent);
- Petroleum & Related Industries - Petroleum - Refineries & Related Industries (7.54 percent); and
- Storage & Transport - Organic Chemical Storage (7.17 percent).

Table III-18. Texas AQS Point Source Benzene Emissions Summary Houston-Galveston Study Area

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	Benzene 2000 Emissions (Peak Ozone Season Day) (tons)	% Total
01	FUEL COMB. ELEC. UTIL.	03	Gas	0.1404	2.39
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0.0062	0.11
02	FUEL COMB. INDUSTRIAL	02	Oil	0.0006	0.01
02	FUEL COMB. INDUSTRIAL	03	Gas	0.6273	10.67
02	FUEL COMB. INDUSTRIAL	04	Other	0.0802	1.36
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	0.0167	0.28
03	FUEL COMB. OTHER	02	Commercial/Institutional Oil	0.0041	0.07
03	FUEL COMB. OTHER	03	Commercial/Institutional Gas	0.0139	0.24
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.0002	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	1.1357	19.31
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	0.024	0.41
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	0.0181	0.31
04	CHEMICAL & ALLIED PRODUCT MFG	04	Agricultural Chemical Mfg	0.0065	0.11
04	CHEMICAL & ALLIED PRODUCT MFG	05	Paint, Varnish, Lacquer, Enamel Mfg	0	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	06	Pharmaceutical Mfg	0.0001	0.00
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	0.6137	10.44
05	METALS PROCESSING	01	Non-Ferrous Metals Processing	0.0011	0.02
05	METALS PROCESSING	02	Ferrous Metals Processing	0.0135	0.23
05	METALS PROCESSING	03	Metals Processing NEC	0.0001	0.00
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	0.0438	0.74
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	0.4434	7.54
06	PETROLEUM & RELATED INDUSTRIES	03	Asphalt Manufacturing	0.0037	0.06
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	0.0181	0.31
07	OTHER INDUSTRIAL PROCESSES	04	Rubber & Miscellaneous Plastic Products	0.009	0.15
07	OTHER INDUSTRIAL PROCESSES	05	Mineral Products	0.0011	0.02
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	0.0114	0.19
07	OTHER INDUSTRIAL PROCESSES	07	Electronic Equipment	0.0004	0.01
07	OTHER INDUSTRIAL PROCESSES	08	Transportation Equipment	0.0008	0.01
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	0.0124	0.21
08	SOLVENT UTILIZATION	01	Degreasing	0.0009	0.02
08	SOLVENT UTILIZATION	02	Graphic Arts	0.0022	0.04
08	SOLVENT UTILIZATION	03	Dry Cleaning	0	0.00
08	SOLVENT UTILIZATION	04	Surface Coating	0.0123	0.21
08	SOLVENT UTILIZATION	05	Other Industrial	0.0013	0.02
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	0.1205	2.05

Table III-18 (continued)

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	Benzene 2000 Emissions (Peak Ozone Season Day) (tons)	% Total
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	1.6672	28.35
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	0.1036	1.76
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	0.4215	7.17
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	0.0334	0.57
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	0.0748	1.27
09	STORAGE & TRANSPORT	10	Inorganic Chemical Transport	0.0003	0.01
09	STORAGE & TRANSPORT	11	Bulk Materials Storage	0.0002	0.00
10	WASTE DISPOSAL & RECYCLING	01	Incineration	0.0257	0.44
10	WASTE DISPOSAL & RECYCLING	03	POTW	0.0218	0.37
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	0.069	1.17
10	WASTE DISPOSAL & RECYCLING	05	TSDF	0.0002	0.00
10	WASTE DISPOSAL & RECYCLING	06	Landfills	0.0734	1.25
10	WASTE DISPOSAL & RECYCLING	07	Other	0.0056	0.10
	Total			5.8804	100.00

These categories are similar to the 2002 draft Point NEI – again noting that Industrial Fuel Combustion – Gas is a significant contributor to this inventory, while it is not a significant contributor to the 2002 draft Point NEI.

It is also noted that the emissions in this inventory are approximately 3 times the emissions in 2002 draft Point NEI if adjusted by multiplying by 365. However, since these annual emissions were calculated based on 365*peak ozone season day, it is expected that these emissions would be significantly overestimated as compared to calculating 365*average day or calculating annual emissions through the application of a temporal profile.

D. POINT SOURCE PROJECTION METHODS

1. *With-CAAA* Projection Calculation

Point source emissions for 2010 *with-CAAA* and 2020 *with-CAAA* were projected based on the 2002 draft NEI point source file using the following algorithm:

Projected emissions= 2002 Emissions(GF)*[1-(CE)]*

where:

GF=Growth Factor for the appropriate source category and year (2010 or 2020)
CE=Control Efficiency (by source category)

Growth factors were developed based on MACT codes and SCCs. Where an emission record contained an associated MACT code, that corresponding growth factor was used in the emission projections. If an emission record had no MACT code, the SCC was used to determine the appropriate growth factor.

Benzene control efficiencies applied to estimate emission reductions associated with post-2002 MACT standards are listed in Table III-19. These control efficiencies are national average values that may over- or under-estimate the emission changes at any individual source. These benzene control efficiencies were taken from the MACT-specific file that was developed by EPA for use with Emissions Modeling System for Hazardous Air Pollutants model (EPA, 2005c).

Table III-19. MACT Specific Control Efficiencies for Post-2002 Standards

MACT Source Category Name	MACT Code	Benzene Control Efficiency
Oil & Natural Gas Production	0501	45.3
Refineries - 10 year	0502	17.15
Refineries - 4 year	0503	62.5
Natural Gas Transmission	0504	45.3
Plastic Parts & Products Surface Coating	0712	77
Carbonyl Sulfide Production the MON	1604	70
Plywood & Composite Work Products	1624	87.6
Pulp & Paper Production Combustion & Non-Combustion (MACT I)	1626	34.2
Pulp & Paper Combustion	1626-1	3.57
Pulp & Paper - Non-combustion	1626-2	35.32
Rubber Tire Manufacturing	1631	70
MON	1640	70

2. *Without-CAAA* Projection Calculations

Point source emissions for 2000, 2010, and 2020 *without-CAAA* scenarios were projected based on the 1990 Point NEI using the following algorithm:

Projected emissions= 2002 Emissions(GF)*

where:

GF=Growth Factor for the appropriate source category and year (2000, 2010 or 2020)

Growth factors were developed based on MACT codes and SCCs. Where an emission record contained an associated MACT code, that growth factor was used. If an emission record has no associated MACT code, the SCC growth factor was applied.

E. EMISSION SUMMARIES

Table III-20 summarizes the *with- and without-CAAA* scenario benzene emission estimates for the study area for each of the scenario years. The following observations can be made from Table III-20:

Table III-20. Houston-Galveston Point Source Benzene Emissions

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	1990 *without-CAAA* (tpy)	2000 *without-CAAA* (tpy)	2000 *with-CAAA* (tpy)	2010 *without-CAAA* (tpy)	2010 *with-CAAA-* (tpy)	2020 *without-CAAA* (tpy)	2020 *with-CAAA* (tpy)
01	FUEL COMB. ELEC. UTIL.	02	Oil	0	0	0	0	0	0	0
01	FUEL COMB. ELEC. UTIL.	03	Gas	0.08	0.12	11.09	0.11	10.5	0.15	14.06
01	FUEL COMB. ELEC. UTIL.	04	Other	0	0	0.33	0	0.35	0	0.35
01	FUEL COMB. ELEC. UTIL.	05	Internal Combustion	0	0	0.42	0	0.4	0	0.53
02	FUEL COMB. INDUSTRIAL	02	Oil	0	0	1.2	0	1.29	0	1.42
02	FUEL COMB. INDUSTRIAL	03	Gas	27.15	30.96	32.45	30.56	35.76	33.86	40.34
02	FUEL COMB. INDUSTRIAL	04	Other	0.03	0.04	11.33	0.05	11.39	0.06	13.24
02	FUEL COMB. INDUSTRIAL	05	Internal Combustion	7.27	8.28	0.63	8.17	0.72	9.08	0.82
03	FUEL COMB. OTHER	04	Misc. Fuel Comb. (Except Residential)	0.01	0.01	0	0.01	0	0.01	0
04	CHEMICAL & ALLIED PRODUCT MFG	01	Organic Chemical Mfg	450.91	542.94	208.23	531.48	245.41	610.03	282.14
04	CHEMICAL & ALLIED PRODUCT MFG	02	Inorganic Chemical Mfg	57.14	68.69	1.99	67.14	2.25	76.97	2.58
04	CHEMICAL & ALLIED PRODUCT MFG	03	Polymer & Resin Mfg	91.99	110.57	0.68	108.08	0.8	123.9	0.93
04	CHEMICAL & ALLIED PRODUCT MFG	04	Agricultural Chemical Mfg	0	0	0	0	0	0	0
04	CHEMICAL & ALLIED PRODUCT MFG	07	Other Chemical Mfg	296.27	354.69	88.85	345.54	106.07	394.32	125.31
05	METALS PROCESSING	01	Non-Ferrous Metals Processing	0	0	0	0	0	0	0
05	METALS PROCESSING	02	Ferrous Metals Processing	0	0	2.06	0	2.65	0	3.52
06	PETROLEUM & RELATED INDUSTRIES	01	Oil & Gas Production	44.7	50.27	10.82	54.49	6.47	50.33	5.74
06	PETROLEUM & RELATED INDUSTRIES	02	Petroleum Refineries & Related Industries	826.07	1051.69	115.27	1187.13	96.77	1359.23	110.79
06	PETROLEUM & RELATED INDUSTRIES	03	Asphalt Manufacturing	0	0	0.28	0	0.29	0	0.35
07	OTHER INDUSTRIAL PROCESSES	03	Wood, Pulp & Paper, & Publishing Products	90.34	102.13	0	133.94	0	170.12	0
07	OTHER INDUSTRIAL PROCESSES	04	Rubber & Miscellaneous Plastic Products	0	0	1.56	0	0.6	0	0.8
07	OTHER INDUSTRIAL PROCESSES	05	Mineral Products	0	0	0.15	0	0.18	0	0.22
07	OTHER INDUSTRIAL PROCESSES	06	Machinery Products	16	19.23	1.92	18.8	2.51	21.55	3.41
07	OTHER INDUSTRIAL PROCESSES	08	Transportation Equipment	8.12	11.38	0	15.15	0	21.11	0

Table III-20 (continued)

Tier 1	Tier 1 Name	Tier 2	Tier 2 Name	1990 *without-CAAA* (tpy)	2000 *without-CAAA* (tpy)	2000 *with-CAAA* (tpy)	2010 *without-CAAA* (tpy)	2010 *with-CAAA-* (tpy)	2020 *without-CAAA* (tpy)	2020 *with-CAAA* (tpy)
07	OTHER INDUSTRIAL PROCESSES	10	Miscellaneous Industrial Processes	61.08	73.42	18.89	71.77	22.61	82.27	26.64
08	SOLVENT UTILIZATION	03	Dry Cleaning	0	0	0.03	0	0.04	0	0.06
08	SOLVENT UTILIZATION	04	Surface Coating	1.63	2.31	0.13	2.76	0.17	3.65	0.22
08	SOLVENT UTILIZATION	05	Other Industrial	0.06	0.08	0.02	0.07	0.03	0.08	0.03
09	STORAGE & TRANSPORT	01	Bulk Terminals & Plants	51.11	65.07	21.96	73.46	26.14	84.1	31.12
09	STORAGE & TRANSPORT	02	Petroleum & Petroleum Product Storage	127.96	155.93	97.41	157.42	73.95	181.38	85.3
09	STORAGE & TRANSPORT	03	Petroleum & Petroleum Product Transport	191.71	256.41	18.48	296.64	21.75	370.47	27.63
09	STORAGE & TRANSPORT	07	Organic Chemical Storage	199.52	239.92	101.44	234.62	120.55	269.06	141.08
09	STORAGE & TRANSPORT	08	Organic Chemical Transport	6.37	7.66	4.83	7.5	5.76	8.62	6.78
09	STORAGE & TRANSPORT	09	Inorganic Chemical Storage	2	2.33	3.42	2.22	4.03	2.45	4.63
10	WASTE DISPOSAL & RECYCLING	03	POTW	7.68	9.1	7.17	10.74	8.3	13.77	10.64
10	WASTE DISPOSAL & RECYCLING	04	Industrial Waste Water	33.51	40.45	9.55	39.68	10.98	45.62	12.89
10	WASTE DISPOSAL & RECYCLING	05	TSDF	4.28	5.08	0.01	5.9	0.01	7.51	0.01
10	WASTE DISPOSAL & RECYCLING	06	Landfills	0	0	0.19	0	0.21	0	0.27
10	WASTE DISPOSAL & RECYCLING	07	Other	0	0	0.09	0	0.11	0	0.13
			Total	**2603**	**3209**	**773**	**3403**	**819**	**3940**	**954**

1. The point source benzene emissions path in the Houston-Galveston study area is dominated by the Federal MACT standards and local VOC control measures that forced the petrochemical facilities in the area to reduce HAP and/or VOC emissions in the years between 1990 and 2000. In Table III-20, significant benzene emission reductions are observed during this period for organic chemical manufacturing, inorganic chemical manufacturing, polymer and resins, and other miscellaneous chemical manufacturing within the chemical industry, and at petroleum refineries.

2. There have been significant reductions in point source benzene emissions in the Houston-Galveston area since 1990 that are attributable to the CAAA. Emissions in 2000 are about 30 percent of those in 1990 (a 70 percent reduction). This overall benzene emission reduction is attributable to emission regulations designed to reduce VOC emissions in this ozone nonattainment area and to CAAA Title III MACT emission standards.

3. Without the CAAA, benzene emissions would have been expected to increase over the study period. For example, point source benzene emissions in 2000 for the *without-CAAA* scenario are 23 percent higher than 1990 levels.

4. Total point source benzene emissions in the Houston-Galveston area are estimated to trend upward between 2000 and 2010, and again between 2010 and 2020 in the *with-CAAA* scenario. The increases are modest – about a 6 percent increase by 2010 and a 16 percent increase between 2010 and 2020. The 2010 to 2020 increase is larger because the analysis is not applying any new VOC or benzene-related control programs post-2010.

5. The point source analysis results are dominated by the organic chemical manufacturing industry and the petroleum refining industry. Organic chemical manufacturing processes, storage, and transport account for 41 percent of the 2002 benzene emissions and 45 percent of the 2010 benzene emissions from point sources. The organic chemical industry was heavily regulated by MACT standards in the 1990 to 2000 time period.

   Refinery benzene emissions are expected to decline between the base year and 2010 as a result of MACT standard effects. After 2010, refinery-emitted benzene is estimated to increase in both the *with-* and *without-CAAA* scenarios as Gulf Coast refinery capacity and utilization is predicted to expand during this period.

6. Where source categories show expected benzene emission decreases between 2002 and 2010 in the *with-CAAA* scenario, those reductions are attributable to post-2002 MACT standard implementation under Title III of the CAAA.

7. In some cases, *with-CAAA* emissions can show a decline in 2010 and then increase in 2020, the most usual reason for this effect would be the interaction between the MACT control efficiencies associated with a source category and the growth factor associated with that category. While the initial control efficiency may be sufficient to reduce emissions, at some point the effect of the activity increase predominates.

8. There is a significant difference between *with*- and *without-CAAA* benzene emissions for the pulp and paper sector, which appears to be attributable to the shutdown of the Champion Paper and Fibre Mill in Pasadena, Texas.

[This page intentionally left blank.]

CHAPTER IV. NONPOINT SOURCE ANALYSIS

Nonpoint emissions were projected for both the *with-CAAA* scenario (2010 and 2020) and the *without-CAAA* scenario (2000, 2010, and 2020). The draft 2002 NEI was used as the initial base for the *with-CAAA* scenario, while the 1990 NEI for HAPs inventory was used as the initial base for the *without-CAAA* scenario. This newly revised inventory (released in November 2005) is available from the EPA. This inventory used as a baseline the 1990 NTI and a number of revisions have been made to it. The primary revisions included converting previously county-level emissions to point source level data wherever feasible. Additional estimates were developed for missing MACT source categories and HAPs so that the baseline inventory was more comparable to the 1999 NEI and 2002 NEI.

A. METHODS FOR THE *WITH-CAAA* SCENARIO

A ranking of nonpoint and nonroad benzene emitting categories (i.e., SCCs) for the 3-county Houston-Galveston area was performed based on benzene emissions reported for EPA's draft 2002 NEI. Table IV-1 presents 2002 SCC-level benzene emissions used to determine priority SCCs for which Pechan focused its efforts for this analysis. As discussed in the sections below, EPA revised 2002 draft NEI estimates for gas marketing to reflect changes for the final NEI, and revised draft NEI commercial marine loading benzene emissions based on updated State-supplied estimates.

1. Gasoline Marketing

Because Gasoline Marketing source categories represent some of the highest nonpoint source benzene emitters in the draft 2002 NEI and because the final 2002 NEI will incorporate major changes to these draft benzene estimates, EPA decided to revise the draft NEI estimates for these categories to approximate the estimates that were incorporated in the final NEI.

In February 2005, EPA notified State and local (S/L) agencies of EPA's intention to replace the draft 2002 NEI Gasoline Marketing VOC and HAP emissions estimates with estimates developed by EPA. The notice requested S/L agency comments on EPA's emission estimation methods and data, and for S/L agencies to identify any instances where EPA's estimates should be replaced with S/L agency estimates. Although a few S/L agencies provided such comments to EPA, no comments were received on the Houston-Galveston area estimates. Therefore, for this study, EPA replaced the draft 2002 NEI estimates with new EPA estimates. These new estimates provide many advantages over EPA-derived estimates from 1999 that were carried forward to the draft 2002 NEI, such as maintaining mass balance for storage and transfer activities such that there is agreement with the activity estimates used for each of the different distribution sectors.

EPA's new nonpoint source emission estimation methodology requires that point source VOC emissions be subtracted from total VOC emissions computed by EPA. Because the final point NEI was not available in time for use in this study, EPA used the draft point NEI in the subtraction calculations. The resulting nonpoint source VOC emissions in each Houston-

Table IV-1. 2002 SCC-Level Houston-Galveston Benzene Nonpoint Source Benzene Emissions

SCC	SCC Description	Emissions (tpy)
2505020000	Storage and Transport ,Petroleum and Petroleum Product Transport, Marine Vessel, Total: All Products	81.84
2501050120	Storage and Transport ,Petroleum and Petroleum Product Storage, Bulk Stations/Terminals: Breathing Loss, Gasoline	42.96
2505040120	Storage and Transport ,Petroleum and Petroleum Product Transport, Pipeline, Gasoline	39.71
2280002100	Mobile Sources ,Marine Vessels, Commercial, Diesel, Port emissions	38.18
2310001000	Industrial Processes ,Oil and Gas Production: SIC 13, All Processes : On-shore, Total: All Processes	36.47
2501080050	Storage and Transport ,Petroleum and Petroleum Product Storage, Airports : Aviation Gasoline, Stage 1: Total	35.13
2810015000	Miscellaneous Area Sources ,Other Combustion, Prescribed Burning for Forest Management, Total	27.72
2104008002	Stationary Source Fuel Combustion ,Residential, Wood, Fireplaces: Insert; non-EPA certified	23.69
2302002200	Industrial Processes ,Food and Kindred Products: SIC 20, Commercial Cooking - Charbroiling, Under-fired Charbroiling	17.57
2104008010	Stationary Source Fuel Combustion ,Residential, Wood, Woodstoves: General	16.94
2610030000	Waste Disposal, Treatment, and Recovery ,Open Burning, Residential, Household Waste (use 26-10-000-xxx for Yard Wastes)	15.80
2810001000	Miscellaneous Area Sources ,Other Combustion, Forest Wildfires, Total	12.54
2401001000	Solvent Utilization ,Surface Coating, Architectural Coatings, Total: All Solvent Types	9.18
2810020000	Miscellaneous Area Sources ,Other Combustion, Prescribed Burning of Rangeland, Total	8.51
2630020000	Waste Disposal, Treatment, and Recovery ,Wastewater Treatment, Public Owned, Total Processed	6.68
2501055120	Storage and Transport : Petroleum and Petroleum Product Storage : Bulk Stations : Breathing Loss : Gasoline	4.86
2501060052	Storage and Transport ,Petroleum and Petroleum Product Storage, Gasoline Service Stations, Stage 1: Splash Filling	3.99
2275050000	Mobile Sources ,Aircraft, General Aviation, Total	3.91
2501060201	Storage and Transport ,Petroleum and Petroleum Product Storage, Gasoline Service Stations, Underground Tank: Breathing and Emptying	3.73
31000299	Industrial Processes ,Oil and Gas Production, Natural Gas Production, Other Not Classified	3.36
2275020000	Mobile Sources ,Aircraft, Commercial Aircraft, Total: All Types	3.27
2275060000	Mobile Sources ,Aircraft, Air Taxi, Total	3.13
2501060100	Storage and Transport ,Petroleum and Petroleum Product Storage, Gasoline Service Stations, Stage 2: Total	3.12
2620030000	Waste Disposal, Treatment, and Recovery ,Landfills, Municipal, Total	3.00
2501060053	Storage and Transport ,Petroleum and Petroleum Product Storage, Gasoline Service Stations, Stage 1: Balanced Submerged Filling	2.89
2302002100	Industrial Processes ,Food and Kindred Products: SIC 20, Commercial Cooking - Charbroiling, Conveyorized Charbroiling	2.46
2501080100	Storage and Transport ,Petroleum and Petroleum Product Storage, Airports : Aviation Gasoline, Stage 2: Total	2.14
2801500000	Miscellaneous Area Sources ,Agriculture Production - Crops, Agricultural Field Burning - whole field set on fire, Total, all crop types	1.08
2810005000	Miscellaneous Area Sources ,Other Combustion, Managed Burning, Slash (Logging Debris), Total	0.49
2280002200	Mobile Sources ,Marine Vessels, Commercial, Diesel, Underway emissions	0.46
2104008004	Stationary Source Fuel Combustion ,Residential, Wood, Fireplaces: Insert; EPA certified; catalytic	0.45
2102006000	Stationary Source Fuel Combustion ,Industrial, Natural Gas, Total: Boilers and IC Engines	0.34
2104008030	Stationary Source Fuel Combustion ,Residential, Wood, Catalytic Woodstoves: General	0.32
2505030120	Storage and Transport ,Petroleum and Petroleum Product Transport, Truck, Gasoline	0.24

Table IV-1 (continued)

SCC	SCC Description	Emissions (tpy)
2285002006	Mobile Sources ,Railroad Equipment, Diesel, Line Haul Locomotives: Class I Operations	0.16
2280003200	Mobile Sources ,Marine Vessels, Commercial, Residual, Underway emissions	0.12
2285002007	Mobile Sources ,Railroad Equipment, Diesel, Line Haul Locomotives: Class II / III Operations	0.12
2660000000	Waste Disposal, Treatment, and Recovery ,Leaking Underground Storage Tanks, Leaking Underground Storage Tanks, Total: All Storage Types	0.07
	All other SCCs < 0.05 tpy	0.20
	Nonpoint (includes nonroad not in NONROAD model) Total Benzene Emissions	**456.81**

Galveston area county were then multiplied by benzene speciation profiles to estimate the nonpoint source benzene emissions in the counties of interest.

Table IV-2 summarizes the benzene emissions by source category for point and nonpoint gasoline marketing activities in the study area.

2. Commercial Marine Loading (SCC 2505020000)

Based on the nonpoint and nonroad SCC-level ranking, commercial marine loading was identified to be the top-emitting SCC for benzene. Volatile organic liquid transfer emissions from ships and barges occur during petroleum product loading and transport. Loading emissions occur when organic vapors in empty cargo tanks are displaced to the atmosphere by the liquid being loaded into the tanks. This SCC represents a composite of loading emissions from all petroleum products, including gasoline, crude oil and refined oil products.

Because of its significance, and because the commercial marine loading emissions were supplied by the State of Texas for the NEI, the TCEQ was contacted to review the estimates for accuracy. Upon review of the estimates, TCEQ provided revised county-level VOC and benzene emissions for this SCC (Rubick, 2005). Benzene emissions were estimated using benzene speciation factors of 0.009 for gasoline and 0.024 for crude oil. These factors were established in an EPA-sponsored 1993 Houston Air Toxics Emissions Inventory (Radian, 1995). Revised total benzene emissions for this source category were estimated to be approximately 82 tpy for the Houston-Galveston area.

Texas rule 115.212 requires at least 90 percent control of VOC emissions from loading and transfer operations at marine terminals. In discussions with TCEQ staff involved with developing the Houston-Galveston State Implementation Plan (SIP), it was determined that this regulation and associated control requirement has been in place since 1993 (Barrett, 2005). As such, the base year 2002 VOC and benzene emissions are assumed to reflect this level of control. No incremental controls were modeled in the future for 2010 and 2020.

3. Bulk Terminals (SCC 2501050120)

One of the priority nonpoint source benzene emission source categories in the Houston-Galveston area is bulk terminals. The nonpoint benzene emission estimates for this source category in the 2002 EPA NEI are from EPA estimates that use information from the EPA Background Information Documents for the proposed emission standards (National Emission Standards for Hazardous Air Pollutants [NESHAPs]) for the gasoline distribution industry (Stage I) to establish national emission levels and average control levels for this source category. Therefore, the benzene (and volatile organic compound [VOC]) emission control level in 2002 represents a combination of the controls in-place across a sample of the bulk terminals in the United States pre-MACT standards as well as the additional HAP and VOC emission reductions achieved by the MACT requirements for bulk terminals. From the information provided in the EPA Background Information Document, the VOC and HAP emissions from bulk terminals have an overall 45 percent control efficiency in 2002 (EPA, 1994a; 1994b). State and local areas with existing bulk terminal VOC requirements are typically those with ozone nonattainment

Table IV-2. Houston-Galveston Benzene Nonpoint Source Benzene Emissions

SCC	Description	Emissions
Point		
40400105	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 10: Breathing Loss (250000 Bbl Capacity)-Fixed Roof Tank	0.0031
40400107	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13: Working Loss (Diam. Independent) - Fixed Roof Tank	0.0001
40400108	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 10: Working Loss (Diameter Independent) - Fixed Roof Tank	0.0001
40400109	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 7: Working Loss (Diameter Independent) - Fixed Roof Tank	0.0206
40400110	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13: Standing Loss (67000 Bbl Capacity)-Floating Roof Tank	0.1676
40400111	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 10: Standing Loss (67000 Bbl Capacity)-Floating Roof Tank	0.1572
40400112	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 7: Standing Loss (67000 Bbl Capacity)- Floating Roof Tank	0.1296
40400115	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 7: Standing Loss (250000 Bbl Cap.) - Floating Roof Tank	0.1964
40400116	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13/10/7: Withdrawal Loss (67000 Bbl Cap.) - Float Rf Tnk	0.6023
40400117	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13/10/7: Withdrawal Loss (250000 Bbl Cap.) - Float Rf Tnk	0.0400
40400148	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13/10/7: Withdrawal Loss - Ext. Float Roof (Pri/Sec Seal)	4.0000
40400151	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Valves, Flanges, and Pumps	1.5796
40400152	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Vapor Collection Losses	0.1035
40400171	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13: Standing Loss - Int. Floating Roof w/ Secondary Seal	0.0183
40400178	Petroleum Liquids Storage (non-Refinery) : Bulk Terminals : Gasoline RVP 13/10/7: Withdrawal Loss - Int. Float Roof (Pri/Sec Seal)	0.0968
40400205	Petroleum Liquids Storage (non-Refinery) : Bulk Plants : Gasoline RVP 10: Working Loss (67000 Bbl. Capacity) - Fixed Roof Tank	0.0143
40400250	Petroleum Liquids Storage (non-Refinery) : Bulk Plants : Loading Racks	0.1913
40400251	Petroleum Liquids Storage (non-Refinery) : Bulk Plants : Valves, Flanges, and Pumps	0.1959
40400252	Petroleum Liquids Storage (non-Refinery) : Bulk Plants : Miscellaneous Losses/Leaks: Vapor Collection Losses	0.0323
40400278	Petroleum Liquids Storage (non-Refinery) : Bulk Plants : Gasoline RVP 10/13/7: Withdrawal Loss - Int. Float Roof (Pri/Sec Seal)	0.0600
40400404	Petroleum Liquids Storage (non-Refinery) : Petroleum Products - Underground Tanks : Gasoline RVP 10: Working Loss	0.0200
40400406	Petroleum Liquids Storage (non-Refinery) : Petroleum Products - Underground Tanks : Gasoline RVP 7: Working Loss	1.2260
40600131	Transportation and Marketing of Petroleum Products : Tank Cars and Trucks : Gasoline: Submerged Loading (Normal Service)	0.7567
40600136	Transportation and Marketing of Petroleum Products : Tank Cars and Trucks : Gasoline: Splash Loading (Normal Service)	0.0000
40600141	Transportation and Marketing of Petroleum Products : Tank Cars and Trucks : Gasoline: Submerged Loading (Balanced Service)	0.3825
40600144	Transportation and Marketing of Petroleum Products : Tank Cars and Trucks : Gasoline: Splash Loading (Balanced Service)	0.0980
NonPoint		
2501050120	Bulk Terminals: Breathing Loss : Gasoline	42.9646
2501055120	Bulk Plants: Breathing Loss, Gasoline	4.8596
2501060052	Gasoline Service Stations : Stage 1: Splash Filling	3.9919
2501060053	Gasoline Service Stations : Stage 1: Balanced Submerged Filling	2.8891
2501060100	Gasoline Service Stations : Stage 2: Total	3.1155
2501060201	Gasoline Service Stations : Underground Tank: Breathing and Emptying	3.7290
2505030120	Truck : Gasoline	0.2420
2505040120	Pipeline : Gasoline	39.7056
	Total	**111.5896**

areas, and emission limits are based on EPA's Control Techniques Guideline for this source category.

(While the average control efficiency at the sources in the Houston-Galveston area may be higher than this, the national average control efficiency is used to compute bulk terminal 2002 emissions in the EPA NEI.)

In future years (2010 and 2020 for the purposes of this study), TCEQ Chapter 115 (115.21-115.217, 115.219) effective November 13, 2003 restricts VOC emissions. For gasoline terminals, the following additional control requirements apply: VOC emissions from the vapor control system vent do not exceed the following rates:

- 0.09 pound per 1,000 gallons of gasoline loaded into transport vessels, plus
- a vapor control system must be used to control the vapors from loading each transport vessel
- each gasoline terminal must be equipped with sensors and other equipment designed and connected to monitor the control device.

It is estimated that meeting Rule 115 requirements will result in an overall VOC (and benzene) control efficiency of 90 percent from uncontrolled emission levels. A benzene control efficiency of 82 percent along with a rule penetration factor of 80 percent was applied to the 2002 emissions for bulk terminals. The 82 percent benzene control efficiency, when applied to a base year emission value with a 45 percent benzene control efficiency, yields a 90 percent control efficiency from uncontrolled emission levels. The rule penetration factor was included to account for the likelihood that there may be bulk terminals that are small enough to not be subject to Rule 115.

4. Pipeline Facilities (SCC 2505040120)

Gasoline distribution pipeline facilities are another one of the priority nonpoint source benzene emission source categories in the study area. Emission estimation methods for the 2002 NEI for this source category parallel those for bulk terminals. Therefore, the benzene (and VOC) emission control level in 2002 represents a combination of the controls in-place across a sample of the pipeline facilities in the United States pre-MACT standards as well as the additional HAP and VOC emission reductions achieved by the MACT requirements for pipeline facilities. From the information provided in the EPA Background Information Document, the VOC and HAP emissions from pipeline facilities have an overall 37 percent control efficiency in 2002 (EPA, 1994a; 1994b).

In 2010 and 2020, TCEQ Chapter 115 (115.211-115.217, 115.219) effective November 13, 2003 restricts VOC emissions from pipeline facilities, to the same level of control as is expected from bulk terminals. Given the current national average level of control for this source category (37 percent), a benzene control efficiency of 84 percent along with a rule penetration factor of 80 percent was applied to the base year emission estimates in order to estimate 2010 and 2020 emissions for pipeline facilities. The 84 percent benzene control efficiency, when applied to a base year emission value with a 37 percent benzene control efficiency, yields a 90 percent control efficiency from uncontrolled emission levels. The rule penetration factor was included to account for the likelihood that there may be pipeline facilities that are small enough to not be affected by Rule 115.

5. Commercial Marine Diesel Engines (SCC 2280002100)

As shown in Table IV-4, in-port commercial marine diesel engines are a significant contributor to Houston-Galveston area benzene emissions. The 2002 base year benzene emissions for in-port commercial marine diesel engines reported in EPA's draft 2002 NEI were based on EPA estimates. However, VOC emissions of 2433 tpy were supplied by the State of Texas for this category for the draft 2002 NEI. Estimation of benzene using heavy-duty vehicle speciation profiles from EPA's NEI (Truex and Norbeck, 1998) applied to the State-supplied VOC emissions yields benzene emissions comparable to, though somewhat higher than the EPA estimates (i.e., 2,433 x 0.0203 = 49 tpy benzene). No adjustments were made to the base year benzene emissions for this category.

Benzene emission reductions are expected from the portion of the diesel vessel fleet estimated to be subject to Category 1 emission standards. EPA has promulgated two sets of commercial marine vessel regulations: a regulation setting Category 1 and 2 marine diesel engine standards, and a regulation setting Category 3 marine diesel engine standards. Category 2 and 3 standards are not estimated to result in VOC emission reductions. Category 1 marine diesel engines are defined as engines of greater than 37 kilowatts, but with a per-cylinder displacement of 5 liters/cylinder or less. VOC emission reductions were estimated using information from the regulatory support documents prepared for this rulemaking (EPA, 1999).

The SCC 2280002100 includes emissions from all engine categories. Using information from the report "Houston-Galveston Area Vessel Emissions Inventory" (Starcrest, 2000), it was estimated that 31 percent of the VOC emissions from in-port diesel vessels are attributable to Category 1 engines (i.e., towboats and other harbor vessels). Emission reductions for 2010 and 2020 were reported to be 12 percent and 24 percent, respectively. When the rule penetration factor described above is accounted for, the overall control factor was estimated to be 4 percent for 2010, and 7 percent for 2020. Benzene emission reductions were estimated to be proportional to these VOC emission changes.

B. METHODS FOR THE *WITHOUT-CAAA* SCENARIO

For consistency with the approach used in the companion national Section 812 study, Pechan revised the 1990 benzene emission estimates for certain source categories as described below.

Table IV-3 displays the SCCs for which Pechan revised the existing 1990 base year inventory emission estimates with estimates derived from backcasting from the 2002 benzene inventory. Because none of these source categories is expected to be affected by control programs that result in benzene emissions reductions between 1990 and 2002, the 1990 emission estimates were generally backcasted by multiplying the 2002 emission estimates by a ratio of 1990 emissions activity to 2002 activity. The two exceptions to this approach were for Prescribed Burning and Forest Wildfires emissions. For these categories, Pechan revised the 1990 emission estimates so that they matched the 2002 estimates. This approach was used to eliminate potentially large fluctuations in emission activity levels that are caused by non-manmade factors (e.g., meteorological conditions). Chapter II of this document provides a discussion of the data sources and methods used in developing 1990 emission activity levels.

Table IV-3. SCCs with Revised Emissions Incorporated into 1990 Inventory

SCC	SCC DESC1	SCC DESC2	SCC DESC3
2104008002	Residential	Wood	Fireplaces: Insert; non-EPA certified
2104008004	Residential	Wood	Fireplaces: Insert; EPA certified; catalytic
2104008010	Residential	Wood	Woodstoves: General
2104008030	Residential	Wood	Catalytic Woodstoves: General
2275020000	Aircraft	Commercial Aircraft	Total: All Types
2275050000	Aircraft	General Aviation	Total
2275060000	Aircraft	Air Taxi	Total
2280002100	Marine Vessels, Commercial	Diesel	Port emissions
2280002200	Marine Vessels, Commercial	Diesel	Underway emissions
2285002006	Railroad Equipment	Diesel	Line Haul Locomotives: Class I Operations
2285002007	Railroad Equipment	Diesel	Line Haul Locomotives: Class II / III Operations
2285002010	Railroad Equipment	Diesel	Yard Locomotives
2810001000	Other Combustion	Forest Wildfires	Total
2810015000	Other Combustion	Prescribed Burning for Forest Management	Total

C. EMISSION SUMMARIES

Table IV-4 summarizes the *with-CAAA* and *without-CAAA* benzene emissions for the Houston-Galveston area for 1990, 2000, 2010, and 2020 for nonpoint sources. Emissions in the nonpoint source analysis *without-CAAA* case are dominated from Synthetic Organic Chemical Manufacturing Industry (SOCMI) fugitives. This is a source category whose emissions are normally reflected in the point source emission inventory, that EPA included in the 1990 NEI for HAPs as a nonpoint source. It is expected that for the 2002 NEI development, SOCMI fugitive emissions were captured in the point source inventory. Therefore, the *with-CAAA* nonpoint source emissions for SOCMI fugitives are zero. Because SOCMI fugitives are such a dominant category in Table IV-4, at the bottom of the table, a without SOCMI fugitive benzene nonpont source emission total is shown as well.

Total Houston-Galveston benzene emissions for this sector are expected to decline slightly (about 4 percent) between 2002 and 2010. This decline is largely attributable to State regulations affecting sources within the Houston-Galveston 1-hour ozone nonattainment area. There are also some expected emission reductions in this period from woodstoves, as certified woodstoves replace non-certified stoves. Prescribed burning and forest wildfire emissions are held constant through the 2002-2020 time horizon.

The most important aggregate source category affecting nonpoint source benzene emissions in the study area in the *with-CAAA* scenario is petroleum and petroleum product storage and transport. This source category accounts for about one-third of the *with-CAAA* nonpoint source benzene emissions in the region. Significant CAAA benzene emission reductions occur for the petroleum storage and transport sector. Many of these emission reductions are attributable to local Houston-Galveston ozone nonattainment area control measures designed to assist the area in meeting the 1-hour ozone NAAQS. These include stringent control requirements for gasoline

Table IV-4. Houston-Galveston Nonpoint Source Benzene Emissions

TIER 2 NAME	TIER 3 NAME	1990	2000 *without CAAA*	2000 *with CAAA**	2010 *without CAAA*	2010 *with CAAA*	2020 *without CAAA*	2020 *with CAAA*
FUEL COMB. ELEC. UTIL.								
Internal Combustion	other	1.78	3.13	0	2.23	0	2.87	0
FUEL COMB. INDUSTRIAL								
Coal	anthracite & lignite	0.00	0.00	0	0.00	0	0.00	0
Coal	other	0.23	0.25	0	0.24	0	0.24	0
Oil	residual	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Oil	distillate	0.00	0.00	0.01	0.00	0.01	0.00	0.01
Gas	natural	0.02	0.03	0.34	0.03	0.35	0.03	0.39
Other	liquid waste	0	0	0.00	0	0.00	0	0.00
Other	other	0	0	0.00	0	0.00	0	0.00
FUEL COMB. OTHER								
Commercial/Institutional Coal	other	5.09	5.50	0.02	5.31	0.01	5.34	0.01
Commercial/Institutional Oil	other	0.00	0.00	0.01	0.00	0.01	0.00	0.01
Commercial/Institutional Gas	other	0.04	0.05	0.05	0.05	0.05	0.06	0.06
Misc. Fuel Comb. (Except Residential)	other	0	0	0.00	0	0.00	0	0.00
Residential Wood	fireplaces	41.84	32.73	24.14	22.86	22.86	22.58	22.58
Residential Wood	woodstoves	33.97	24.43	17.26	16.44	16.44	15.83	15.83
Residential Other	distillate oil	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Residential Other	natural gas	0.04	0.03	0.04	0.04	0.05	0.04	0.05
Residential Other	bituminous/subbituminous coal	7.34	4.86	0.00	6.00	0.00	5.44	0.00
Residential Other	other	0	0	0.00	0	0.00	0	0.00
CHEMICAL & ALLIED PRODUCT MFG								
Organic Chemical Mfg	socmi fugitives	2371.43	2850.46	0	2786.43	0	3194.32	0
Organic Chemical Mfg	other	12.15	15.62	0	16.06	0	19.22	0
METALS PROCESSING								
Metals Processing NEC	other	0.28	0.35	0	0.41	0	0.54	0
PETROLEUM & RELATED INDUSTRIES								
Oil & Gas Production	natural gas	27.95	32.57	3.36	33.05	3.47	36.21	3.21
Oil & Gas Production	other	37.68	33.18	36.47	35.87	34.50	31.84	31.25
OTHER INDUSTRIAL PROCESSES								
Agriculture, Food, & Kindred Products	commercial cooking	0	0	20.03	0	23.92	0	30.35
Mineral Products	other	0	0	0.01	0	0.01	0	0.01
SOLVENT UTILIZATION								
Surface Coating	architectural	5.45	6.57	9.18	7.43	10.11	8.30	11.27
Other Industrial	other	0.03	0.04	0	0.04	0	0.04	0
Nonindustrial	pesticide application	0	0	0.00	0	0.00	0	0.00
Nonindustrial	adhesives	0	0	0.00	0	0.00	0	0.00
Nonindustrial	consumer solvents	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table IV-4 (continued).

TIER 2 NAME	**TIER 3 NAME**	**1990**	**2000 *without-CAAA***	**2000 *with-CAAA****	**2010 *without-CAAA***	**2010 *with-CAAA***	**2020 *without-CAAA***	**2020 *with-CAAA***
STORAGE & TRANSPORT								
Bulk Terminals & Plants	area source: gasoline	59.21	73.63	47.82	90.67	23.10	108.94	27.75
Petroleum & Petroleum Product Storage	other	4.48	5.56	37.26	6.79	43.74	8.12	52.25
Petroleum & Petroleum Product Transport	other	49.30	62.85	121.79	71.14	111.03	81.45	128.00
Service Stations: Stage I	other	19.45	24.19	6.88	29.78	8.09	35.78	9.72
Service Stations: Stage II	other	18.84	23.43	3.12	28.85	3.66	34.66	4.40
Service Stations: Breathing & Emptying	other	0	0	3.73	0	4.39	0	5.27
WASTE DISPOSAL & RECYCLING								
Open Burning	residential	13.55	17.07	15.80	20.11	17.96	23.26	20.78
POTW	wastewater treatment	1.07	1.27	6.68	1.50	7.74	1.92	9.92
Landfills	industrial	0.00	0.00	0	0.00	0	0.00	0
Landfills	other	2.13	2.52	3.00	2.98	3.47	3.82	4.45
Other	other	0	0	0.07	0	0.07	0	0.08
OFF-HIGHWAY								
Aircraft	other	9.16	10.84	10.30	11.92	11.92	14.38	14.38
Marine Vessels	diesel	42.79	51.06	38.77	59.04	41.43	66.51	42.82
Railroads	other	0.23	0.26	0.33	0.36	0.36	0.38	0.38
MISCELLANEOUS								
Other Combustion	agricultural fires	0	0	1.08	0	1.11	0	1.17
Other Combustion	prescribed burning	27.72	27.72	27.22	27.72	27.22	27.72	27.22
Other Combustion	forest wildfires	12.54	12.54	12.54	12.54	12.54	12.54	12.54
Other Combustion	slash burning (logging)	0	0	0.49	0	0.56	0	0.66
Other Combustion	other burning	0	0	8.5	0	8.5	0	8.5
TOTALS		2,806	3,323	457	3,296	439	3,762	486
Totals without SOCMI Fugitives		435	473	457	510	439	568	486

* - Values are from 2002 NEI, but are used to represent 2000 emissions

storage tanks, transfer operations at bulk terminals, and Stage I (filling underground tank) and Stage II (refueling vapor recovery) controls at service stations. Because many of the control programs that reduce VOC and benzene from petroleum storage and transport were instituted by 2000, the post-2000 benzene emissions from some source categories within this larger category can be higher than the zero estimates.

Total nonpoint source benzene emissions in the area are estimated to increase between 2010 and 2020. While there may be regulations added in this area in the next few years to meet new nonattainment obligations, based on the current set of Federal and State regulations affecting this area, nonpoint source sector categories' benzene emission rates have no expected declines in the 2010 to 2020 period other than for woodstoves. The 2020 *with-CAAA* nonpoint source benzene emissions are estimated to be 10 percent higher than the 2010 levels.

In the *without-CAAA* scenarios there is a steady increase in benzene nonpoint source emissions over the study period. Total emissions in this scenario increase by 34 percent from 1990 to 2020.

There are some source categories whose benzene emissions were included in the 2002 NEI, but not in the 1990 NEI. Commercial cooking is one example (with 20 tpy benzene) in the 2000 *with-CAAA* scenario. Commercial cooking is a source type whose emission rates were not measured until recently.

[This page intentionally left blank.]

CHAPTER V. ON-ROAD VEHICLE ANALYSIS

A. METHODS

1. Overall Approach

On-road emissions were calculated for the 8-county Houston area. The ultimate purpose of this emission inventory was for use in the AERMOD modeling system. Therefore, it was determined that link-level (i.e., roadway segment) emissions data by hour of day for each season would be the most appropriate level of detail. VMT data were prepared at this level of detail based on VMT data files prepared for the Houston area by the Houston-Galveston Area Council and further processed by the Texas Transportation Institute (TTI). The most current VMT data analysis prepared by these agencies during the summer of 2005 was used for this modeling. However, as discussed below, adjustments were needed to apply these data to the temporal needs of the Section 812 study. Emission factors were calculated using EPA's MOBILE6.2 model. Where possible, local input data for the Houston area, as provided by TCEQ, were used in the development of the MOBILE6.2 input files. The sections below describe the adjustments made to the VMT data files and the preparation of the MOBILE6.2 input files.

The on-vehicle emissions modeling approach includes technological changes that would affect evaporative emissions, such as lower volatility fuel, new evaporative test procedures, onboard vapor recovery system emission controls, and reductions from I/M programs.

2. VMT

TTI provided link-level hourly VMT data for the years 2002, 2009, and 2012. The data included the eight counties in the Houston nonattainment area. Because the data were developed for ozone modeling, the link-level data files had been developed specific to an August/September episode. The data were supplied for four day types: weekday, Friday, Saturday, and Sunday. Several adjustments were made to the VMT data to make it appropriate for use in the Section 812 benzene study.

First, the VMT data were adjusted from the specific modeling period to the four seasons. TTI provided adjustment factors to be used with the data set to adjust the VMT from the episode-specific time period to each of the four seasons. These seasonal adjustment factors were provided for each of the four day types, but did not differ by year and are shown in Table V-1. Multiplying these adjustment factors by the VMT in the episode-specific VMT data sets, four new seasonal data sets were prepared from each of the episode-specific VMT data sets for 2002, 2009 and 2012. These new data sets were still at the link and hourly level of detail.

Next, the VMT data were allocated by vehicle type. TTI provided VMT fractions by vehicle type corresponding to each data set. The VMT fractions differed by time of day (AM peak, PM peak, overnight, and midday) and road class (freeways, arterials, and collectors). These data were matched appropriately to the hour and functional class of each link VMT and then each hourly, link-level VMT record was multiplied by each of the corresponding VMT mix fractions to break the link-level VMT data out by vehicle type.

Table V-1. Seasonal and Day of Week VMT Adjustment Factors

Season	Day Type	VMT Adjustment Factor
Summer	Weekday	1.00084
Summer	Friday	0.96223
Summer	Saturday	0.95763
Summer	Sunday	1.00927
Fall	Weekday	1.00524
Fall	Friday	0.96856
Fall	Saturday	0.96464
Fall	Sunday	0.98878
Winter	Weekday	0.97610
Winter	Friday	0.94522
Winter	Saturday	0.92773
Winter	Sunday	0.94283
Spring	Weekday	1.01947
Spring	Friday	0.99227
Spring	Saturday	1.00435
Spring	Sunday	1.05806

After the VMT data were allocated by vehicle type, the seasonal VMT data sets were adjusted to combine the VMT from the four day types into a single VMT value representing an average day of the week. This was done by weighting the four different day type VMT values from a given link, hour, vehicle type, and season by the number of that type of day that occurs in that season and year. For example, in winter 2002 there were 53 weekdays (excluding Fridays), 12 Fridays, 12 Saturdays, and 13 Sundays, for a total of 90 winter days. Thus, the following equation was used to calculate the VMT from a given link, vehicle type, and hour in winter 2002 for an average day:

$$VMT(avgday) = (VMT(wkd)*53 + VMT(Fri)*12 + VMT(Sat)*12 + VMT(Sun)*13)/90$$

The hourly, link-level speed values were similarly weighted to obtain hourly, link-level speeds for an average day.

The final step in preparing VMT that could be used to calculate on-road emissions for Houston was to adjust the VMT to the study years. To obtain the 2000 VMT, the 2002 and 2009 link-level VMT data were matched by link, hour, vehicle type, and season. The average annual growth rate from 2002 to 2009 was then calculated from each set of matching records. The 2002 VMT data was then extrapolated back by two years using this average annual growth rate to estimate 2000 VMT. The link network from the 2002 data set was assumed to be the same as the 2000 link network. Thus, links included in the 2009 data file, but not in the 2002 data file were not included in the 2000 data file. Similarly, the 2010 VMT was calculated using the 2009 and 2012 link-level data files. Again, the average annual growth rate was calculated between these two years for the VMT values matched on each link, hour, vehicle type, and season. One year of this average annual growth rate was applied to the 2009 link-level VMT data set to obtain a 2010 link-level VMT data set. The 2010 link network was assumed to be the same as the 2009 link network. The 2020 VMT data set was calculated based on extrapolating the 2002 and 2012 VMT data. The average annual growth rate was calculated between these two years for the

VMT values matched on each link, hour, vehicle type, and season. Eight year of this average annual growth rate was applied to the 2012 link-level VMT data set to obtain a 2020 link-level VMT data set. The 2012 link network was used in 2020.

VMT data for 1990 were not available at the link level. However, weekday (Monday through Friday) VMT summaries for the 8-county Houston area were provided by TCEQ, based on data included in rate-of-progress plans. A comparable data summary was prepared from the 2002 episode-specific VMT data, after weighting the 2002 weekday VMT (multiplied by 4/5) and the 2002 Friday VMT (multiplied by 1/5) to estimate 2002 average weekday VMT. Backcasting factors were then calculated from 2002 to 1990 by county and functional road class (e.g., 1990 VMT divided by 2002 VMT). These backcasting factors were then matched to the 2002 episode-specific VMT data and multiplied by the corresponding 2002 VMT values to obtain a 1990 episode-specific hourly, link-level VMT file for each of the four day types in the 2002 link-level VMT file. From this point, the adjustments applied to the 1990 VMT were similar to those applied in the other years. First, the VMT were allocated to the four seasons, using the same seasonal VMT adjustment factors listed in Table V-1. Next, the VMT were allocated by vehicle type, using 1990 vehicle mix fractions by time of day and facility type, specific to 1990, which were also provided by TCEQ. These vehicle-type specific VMT data were then aggregated to obtain an average day VMT for each hour, link, season, and vehicle type. The 2002 link network was also used for 1990.

It should be noted that in this analysis, speeds have been calculated by weighting them by VMT when the VMT data were adjusted. If resources permitted, it would have been preferable to recalculate the link speeds based on the VMT volumes in the seasonal, average day VMT files.

3. Emission Factor Inputs

a. Registration Distributions

The default MOBILE6.2 registration distribution of vehicles by age was used in 1990. In 2000, for both the *with-CAAA* and the *without-CAAA* scenarios, three different registration distribution files were used. One file was specific to Harris County. Brazoria, Fort Bend, Galveston, and Montgomery Counties were all modeled with the same urban 2002 registration data file. Chambers, Liberty, and Waller Counties were modeled with a 2002 rural registration distribution file. In both 2010 scenarios, the registration distributions varied for each county. All of these registration distribution files were downloaded from TCEQ’s web site. The 2010 registration distributions were also used in 2020.

b. Diesel Sales Fractions

The default MOBILE6.2 diesel sales fractions were used in 1990. For both scenarios in 2000, Statewide diesel sales fraction estimates were used in all counties. Similarly in 2010 and 2020, Statewide diesel sales fraction estimates were applied in all counties in both scenarios. The diesel sales data were available from TCEQ’s web site.

c. Inspection and Maintenance (I/M) and Anti-Tampering Program (ATP) Inputs

An I/M program was not present in the Houston area in 1990. However, Harris County did have an ATP implemented in 1990. Therefore, only the Harris County ATP was modeled in 1990. Since the *without-CAAA* scenarios for 2000 and 2010 represent emissions without CAAA controls, only the Harris County ATP was modeled in those scenarios, with no I/M programs in any county. In the 2000 *with-CAAA* scenario, an I/M program was modeled in all eight Houston area counties. However, the program inputs for Harris County differed from those of the remaining 7 counties. In addition, only Harris County was modeled with an ATP in the 2000 *with-CAAA* scenario. By 2010 in the *with-CAAA* scenario, all eight counties were modeled with both I/M programs and ATPs. However, the program inputs varied for Harris County, the urban counties, and the rural counties. Again, all I/M program data were available from the TCEQ web site. The 2020 data used were the same as the 2010 data.

d. Temperature Inputs

Since hourly emissions were the desired end product, hourly temperature input data were used. For both 2000 scenarios, hourly temperature and humidity inputs by month were used. These were obtained from the national county database associated with EPA's NMIM. For the 2010 and 2020 scenarios, long-term average hourly temperature and humidity inputs for each month were obtained from the same source. Both the 2000 and long-term average temperature and humidity data sets varied by county. For 1990, an hourly temperature data set for Harris County from the TCEQ web site was used to develop 1990 hourly temperature inputs by month. This same data set was applied to all counties. The MOBILE6.2 default humidity inputs were used in 1990.

e. Fuel Inputs

Monthly fuel inputs, based on 2000 fuel data, from the NMIM national county database were used as the basis of the 2000 and 2010 *with-CAAA* fuel inputs. The 2010 and 2020 fuel data were the same as the 2000 fuel data, but with gasoline sulfur levels adjusted to account for the low sulfur gasoline provisions that began taking effect in 2004. For 1990, and the 2000 and 2010 *without-CAAA* scenarios, July and January 1990 historical fuel data, used in EPA's 1990 NEI, were applied to all eight counties. The July fuel data was applied in months May through September, while the January data were applied in the other (winter season) months. The fuel inputs for all years included Reid vapor pressure (RVP), as well as information on the oxygen content of the fuel, the benzene content, and additional properties needed to model toxic air pollutants in MOBILE6.2.

f. Development of MOBILE6.2 Input Files

Using the inputs discussed above, MOBILE6.2 input files were prepared for 1990, and the *without-CAAA* and *with-CAAA* 2000, 2010, and 2020 scenarios. MOBILE6.2 scenarios were developed by hour, using the hourly temperature and humidity inputs. In all years and scenarios, a set of emission factors were developed for the MOBILE6.2 freeway and arterial road categories in 5 mile per hour increments. Additionally, a single scenario representing ramps and one representing local roads were also modeled. (Emission factors for these two road categories do not vary by speed in MOBILE6.2.) In addition to modeling the 1990 fuel characteristics and the 1990 Harris County ATP, the 2000, 2010, and 2020 No CAAA input files included the "NO CAAA" command to exclude the effects of national CAAA programs on the emission factors. For the *with-CAAA* scenarios, this command was removed and the actual I/M programs and ATPs discussed above, and the fuel programs appropriate to the year being modeled were included.

B. EMISSION SUMMARIES

Table V-2 summarizes the 8-county benzene emission totals by year and scenario for each vehicle type. For comparison, emissions from the draft EPA 2002 NEI are also included in this table. While this table shows that some benzene emission reductions would occur through 2010 without the additional provisions of the CAAA, the reductions obtained through the CAAA control measures are significant. From 1990, with the CAAA programs in place, benzene emissions in Houston are reduced by 65 percent in 2000, by 85 percent in 2010, and by 86 percent in 2020. While the reductions from 1990 to 2020 with the CAAA programs in place are similar to the reductions achieved between 1990 and 2010, an 86 percent reduction in benzene emissions is achieved between the *with-CAAA* and the *without-CAAA* scenarios in 2020, compared to a 77 percent reduction between emissions in these two scenarios in 2010. Thus, despite a significant increase in VMT from 2010 to 2020, benzene emissions continue on a downward trend through 2020 in the *with-CAAA* scenario.

MOBILE6.2 has a "NO CAAA" command which enables a user to model emissions without reductions under the Clean Air Act Amendments of 1990. However, results using this command still include reductions from controls in place prior to the 1990 amendments. Based on an interim set of runs for 2000 and 2010 that were performed with the MOBILE6.2 "NO CAAA" command, but with the fuel and I/M programs and ATP specific to either 2000 or 2010, it was determined that in 2000, about 96 percent of the benzene emission reductions that occur between the 2000 No CAAA and the 2000 *with-CAAA* scenarios are attributable to changes in fuel program parameters and the I/M program, with the remaining reductions due to the phasing in of more stringent emission standards under the 1990 Clean Air Act Amendments. By 2010, a greater portion of the emission reductions (29 percent) is attributable to greater phase-in of new emission standards.

Table V-2. Houston 8-County Annual Benzene Emissions (tpy)

Vehicle	1990	2002 NEI	2000 *without-CAAA*	2000 *with-CAAA*	2010 *without-CAAA*	2010 *with-CAAA*	2020 *without-CAAA*	2020 *with-CAAA*
LDGV	1,528.96	816.54	1,073.67	530.39	1,031.74	225.02	1,509.62	199.12
LDGT1	738.81	255.97	614.41	316.23	593.70	130.26	866.35	121.61
LDGT2	236.76	104.10	108.30	57.92	131.82	29.16	214.11	34.83
HDGV	133.06	25.02	41.85	21.44	21.24	7.95	24.78	5.34
MC	2.27	2.15	3.48	1.94	3.19	1.94	4.24	2.57
LDDV	3.27	0.68	0.58	0.57	0.20	0.09	0.16	0.03
LDDT	4.40	0.91	0.73	0.67	0.66	0.28	0.81	0.15

HDDV	78.06	27.39	26.25	23.57	23.38	15.58	32.70	14.26
Total	**2,725.59**	**1,232.76**	**1,869.27**	**952.75**	**1,805.92**	**410.28**	**2,652.77**	**377.92**

Table V-3 summarizes the annual onroad benzene emissions from the 3-county Houston area, including Brazoria, Galveston, and Harris Counties. This table was prepared to enable consistent comparisons with the other emission sectors where emissions were only calculated for these three counties. Emissions from these three counties account for about 87 percent of the 8-county emission totals in 1990, about 80 percent of the 8-county total by 2010, and about 75 percent of the 8-county total by 2020. The benzene on-road emission percentage reductions for these three counties from the *without-CAAA* scenarios to the *with-CAAA* scenarios are similar to those observed for the 8-county region. However, the percentage reductions in benzene emissions in the *with-CAAA* scenario from 2010 to 2020 are greater for the 3-county region (about 14 percent) compared to those for the 8-county region (about 8 percent). This difference is primarily driven by the greater rate of VMT increase for the 8-county region as a whole compared to the 3-county region, implying that VMT growth is occurring at higher rates outside of the core metropolitan area. Table V-4 summarizes the annual VMT in each of the analysis years for the 8-county region and the 3-county region.

Table V-3. Annual On-road Benzene Emissions from Brazoria, Galveston, and Harris Counties (emissions in tpy)

Vehicle	1990	2000 *without-CAAA*	2000 *with-CAAA*	2010 *without-CAAA*	2010 *with-CAAA*	2020 *without-CAAA*	2020 *with-CAAA*
LDGV	1,335.09	889.04	421.80	831.73	180.82	1,133.47	148.76
LDGT1	642.58	504.70	254.44	474.18	103.16	645.71	90.05
LDGT2	205.73	88.13	46.27	104.57	22.96	160.55	25.91
HDGV	115.44	33.71	16.95	16.94	6.34	18.90	4.06
MC	1.95	2.82	1.57	2.53	1.54	3.25	1.97
LDDV	2.85	0.48	0.47	0.16	0.07	0.13	0.02
LDDT	3.83	0.60	0.55	0.52	0.22	0.62	0.12
HDDV2b	1.87	1.20	0.78	1.81	0.69	2.87	0.83
LHDDV	2.94	1.72	1.22	1.93	0.87	2.56	0.76
MHDDV	4.97	2.68	2.43	2.57	1.61	3.35	1.26
HHDDV	57.17	15.58	14.68	11.38	8.74	14.50	7.37
BUS	1.15	1.01	0.88	1.09	0.63	1.66	0.66
Total	**2,375.58**	**1,541.66**	**762.04**	**1,449.40**	**327.64**	**1,987.58**	**281.78**

Table V-4. Annual Vehicle Miles Traveled (in million miles)

Area	1990 VMT	2000 VMT	2010 VMT	2020 VMT
8-county Houston Area	36,029	44,346	53,061	78,182
Brazoria, Galveston, and Harris Counties	31,059	36,453	42,355	58,436

It should be noted that EPA has recently determined that cold start HC emissions have been underestimated in MOBILE6 at temperatures below 75 degrees Fahrenheit, based on engine certification data collected primarily from Tier 1 and more recent vehicles to comply with the

CO cold temperature standards. This would likely lead to a similar underestimation of benzene at cold temperatures. As a result, the benzene estimates from this analysis are likely to have been underestimated in the *with-CAAA* scenario, though the effect in the Houston area would not be as significant as it would be in colder areas of the country. Modeling for EPA's Mobile Source Air Toxics (MSAT) rule, finalized in February 2007, indicates the underestimate for Houston in 2010 would be about 10%. These results suggest that MOBILE6 is likely to introduce a minor upward bias on mobile source benzene emission reductions due to the CAAA.

Of greater significance, however, is the exclusion of the MSAT rule requirements from this analysis. The MSAT rule, which was still in development at the time this analysis was performed, includes a limit on the benzene content of gasoline and a cold temperature hydrocarbon standard which would further reduce benzene emissions from the with-CAAA scenario, and thus increase the benefits of CAAA benzene controls in 2020. However, these reductions were not accounted for in this analysis, and therefore estimate changes in mobile source benzene emissions between the *with*- and *without-CAAA* scenarios are most likely underestimated.

[This page intentionally left blank.]

CHAPTER VI. NONROAD ENGINE/VEHICLE ANALYSIS

This report section describes the base and projection year emission estimation methods and results for the nonroad engine/vehicle sector.

A. METHODS

Nonroad benzene emission estimates for the Houston, Texas area were developed by multiplying benzene speciation factors by VOC emissions output from EPA's NONROAD2004 model. The benzene speciation factors and fuel data inputs were obtained from EPA's National Mobile Inventory Model (NMIM), which was used to develop the final 2002 NEI nonroad estimates. The 2002 NEI was not used, however, to represent 2000 year nonroad benzene emissions, since comparable estimates were not available for the other analysis years. NONROAD model equipment categories include: recreational vehicles, farm and construction machinery, lawn and garden equipment, aircraft and rail support equipment, and other industrial and commercial applications. Nonroad equipment categories not modeled by NONROAD, (i.e., aircraft, commercial marine and locomotives), are included in the nonpoint area source portion of the inventory, and methods for these categories are addressed in Chapter II. The NONROAD2004 model was released by EPA's Office of Transportation and Quality (OTAQ) in May 2004 (EPA, 2004a). This version of the model incorporates all Federal engine standards, with the exception of the large spark-ignition evaporative standards. VOC reductions from this standard were applied outside of the NONROAD model.

In the main Section 812 Prospective study, a revised NONROAD model growth file was prepared to model region-specific growth rates for all regions of the United States. The procedures used to develop the regional growth rates are described in the Section 812 Prospective report (Pechan, 2005a). This region-specific NONROAD growth file was used for the Houston-area model runs as well. Input files were prepared for Brazoria, Galveston, and Harris counties to reflect the appropriate temperature and fuel inputs for the *with-CAAA* scenario runs. In addition, fleet emission rate inputs were modified to remove the effect of CAAA-related standards for the *without-CAAA* runs. Using county-specific input files, NONROAD model runs were performed to generate seasonal emission estimates for each scenario year. Seasonal emissions were then summed to estimate annual emissions at the county and SCC level for each scenario/year.

The NONROAD model accounts for technological changes in the population of nonroad engines, retiring older (higher-emitting) engines and assuming that new engines sold will meet the federal emission control mandates applicable in the year of sale. The NONROAD model does not account for changes in use (e.g., average annual hours of use by engine type) during the projection years; average annual hours of use are assumed to stay constant at historical levels throughout the forecast period.

1. NONROAD Model Inputs

The NONROAD model uses ASCII format input files, termed *option* files, that specify the parameters for a specific model run, including ambient temperature and fuel characteristics for the modeled geographic area and time period. The appropriate temperature and fuel data inputs were compiled for each of the years of interest (1990, 2000, 2010, and 2020).

Fuel specifications for *without-CAAA* scenario runs for 2000 and the forecast years were based on 1990 data. Statewide average temperatures were used for 1990 and seasonal average county-level temperatures were compiled for 2000. The 2010 and 2020 temperatures were based on 30-year historical temperature trends. Year-specific temperatures for each year and county are shown in Table VI-1 (EPA, 2004b). RVP and fuel sulfur levels for each season are shown in Table VI-2 (these inputs were the same for all three counties). Using these data, seasonal, county-specific NONROAD model option files were developed for each of the three Texas counties for 1990, 2000, 2010, and 2020.

Table VI-1. Year-Specific Temperatures

County	Season	Minimum	Maximum	Average
1990 Temperature, degrees Fahrenheit				
All Counties	Winter	46	67	57
	Spring	61	79	70
	Summer	73	95	84
	Autumn	59	82	71
2000 Temperature, degrees Fahrenheit				
Brazoria	Winter	48	66	56
	Spring	62	80	70
	Summer	73	92	82
	Autumn	61	79	69
Galveston	Winter	46	66	55
	Spring	64	79	71
	Summer	75	92	83
	Autumn	61	79	69
Harris	Winter	46	67	55
	Spring	62	81	71
	Summer	74	95	84
	Autumn	61	80	69
2010 and 2020 Temperature, degrees Fahrenheit				
Brazoria	Winter	46	65	54
	Spring	60	78	68
	Summer	74	91	82
	Autumn	61	81	70
Galveston	Winter	47	64	54
	Spring	62	77	69
	Summer	76	90	83
	Autumn	63	80	71
Harris	Winter	43	64	52
	Spring	59	79	68
	Summer	73	93	83
	Autumn	60	81	69

Table VI-2. Seasonal Fuel Input Values for Brazoria, Galveston, and Harris Counties for *Without-CAAA* Scenario[1]

Season	RVP, psi	Gasoline Sulfur, ppm	Diesel Sulfur, ppm
Winter	12.3	339	2500
Spring	9.8	339	2500
Summer	8.1	339	2500
Autumn	8.7	339	2500

NOTE: [1]Applies to 1990, 2000 *without*-, 2010 *without*-, and 2020 *without-CAAA* model runs.

For modeling a *without-CAAA* scenario for all years, the NONROAD model technology file (TECH.DAT) was revised to remove the effect of Federal nonroad control programs. A base technology with a sales fraction of 1 is reported for each equipment category or application for the year 1900. This fraction changes to account for sales of engines equipped with technologies needed to meet a specific standard for each implementation year. To ensure that the model applies emission rates corresponding to the *base* technology type for all *without-CAAA* runs, sales fractions for all other years besides 1900 were removed in the TECH.DAT file. The one exception to this was the T0 technology type for diesel engines. This technology type, which applies to engines sold in 1988 and later, was retained in the TECH.DAT files, since this occurred prior to the CAAA, and was a result of the "spillover" of highway diesel control technology.

Input parameters for Brazoria, Galveston, and Harris counties were developed that reflected local and national fuel programs for the *with-CAAA* scenario runs for 2000, 2010, and 2020. Local inputs, including seasonal RVP limits, oxygenated fuel specifications for reformulated gasoline, and Stage II programs were available from EPA's NMIM county database (EPA, 2005d). Federal gasoline and diesel fuel sulfur levels were incorporated as well. Tables VI-3a and VI-3b list the seasonal fuel parameter inputs used for all three counties for each modeling year.

Table VI-3a. Seasonal Reformulated Gasoline Characteristics for Brazoria, Galveston, and Harris Counties[1]

Season	O_2 Weight %	RVP, psi
Winter	1.91	11.9
Spring	1.96	8.9
Summer	2.00	6.7
Autumn	1.99	7.4

NOTE: [1]Applies to 2000, 2010, and 2020 *with-CAAA* model runs

Table VI-3b. Seasonal Fuel Sulfur Content for Brazoria, Galveston, and Harris Counties

Year	Season	Gasoline Sulfur, ppm	Diesel Sulfur, ppm
2000	Winter	219.0	2284
	Spring	180.1	2284
	Summer	151.0	2284
	Autumn	160.7	2284
2010	Winter	30	165
	Spring	30	165
	Summer	30	165
	Autumn	30	165
2020	Winter	30	11
	Spring	30	11
	Summer	30	11
	Autumn	30	11

The only Federal standards not modeled by NONROAD2004 include permeation and evaporative emission standards for gasoline recreational and large spark-ignition (S-I) engines, respectively. The evaporative standards for recreational equipment only affect permeation emissions, which are not currently included in NONROAD2004. These standards do not affect any other evaporative emission components in the model (i.e., diurnal or refueling). Therefore, recreational equipment permeation emission standards benefits were not included in the *with CAAA* scenario.

For the large S-I evaporative standard, base and control case future year inventories compiled by EPA were used to calculate emission reductions in 2010 and 2020 (EPA, 2002). These emission reductions vary by evaporative component, but for this analysis, emissions were summed across all evaporative components to estimate emission reductions. Large S-I evaporative benzene emission reductions were estimated to be 59.7 percent in 2010 and 82 percent in 2020.

A rule penetration adjustment was calculated to account for the fraction of the SCC-level emissions that are affected by the rule. Since the rule only affects S-I engines greater than 25 horsepower, an adjustment was developed to reflect that fraction of the activity associated with these larger engines. This was estimated using 2002 national gasoline consumption results by horsepower and equipment category from NONROAD2004. As a simplifying assumption, the 2002 rule penetration value was used for 2010 and 2020 and for all applications within a category, though this is likely to vary by year and application. Table VI-4 provides a summary of the horsepower-related rule penetration values by equipment category. Table VI-5 provides the overall control effectiveness or emission reductions estimated for each nonroad category for 2010 and 2020. These emission reductions were applied directly to the SCC-level VOC evaporative emissions output from the NONROAD model as a post-processing step.

Table VI-4. Horsepower-Related Rule Penetration Values by Category for Large S-I Evaporative Standards

Fuel Type	Classification	Rule Penetration, %
Gasoline	Agricultural Equipment	40
	Airport Equipment	74
	Commercial Equipment	5
	Construction and Mining Equipment	14
	Industrial Equipment	59
	Commercial Lawn and Garden Equipment	7
	Railroad Equipment	4
	Recreational Equipment[1]	43
Compressed Natural Gas (CNG)	All Classifications	100
Liquefied Petroleum Gas (LPG)	All Classifications	100

NOTE: [1]Applies to specialty vehicle carts only; other recreational equipment covered by recreational standards.

Table VI-5. Control Effectiveness Values by Year and Category for Large S-I Evaporative Standards

Fuel Type	Classification	2010 Control Effectiveness, %	2020 Control Effectiveness, %
Gasoline	Agricultural Equipment	23.6	32.4
	Airport Equipment	43.9	60.3
	Commercial Equipment	3.0	4.2
	Construction and Mining Equipment	8.1	11.1
	Industrial Equipment	35.3	48.5
	Commercial Lawn and Garden Equipment	4.1	5.6
	Railroad Equipment	2.4	3.3
	Recreational Equipment[1]	25.5	35.0
CNG	All Classifications	59.7	82.0
LPG	All Classifications	59.7	82.0

NOTE: [1]Applies to specialty vehicle carts only; other recreational equipment covered by recreational standards.

The following equation provides an example of how overall adjusted emission reductions were estimated for 4-stroke gasoline industrial equipment in 2010:

$$ER_{ADJ} = RP_{hp} \times ER$$

where:

ER_{ADJ} = adjusted emission reduction accounting for rule penetration
RP_{hp} = rule penetration for affected horsepower fraction (from Table VI-4)
ER = evaporative emission reduction for affected engines (from Table VI-5)

$$
\begin{aligned}
ER_{ADJ} &= 0.590 \times 0.5974 \\
&= 0.353 \\
&= 35.3 \text{ percent}
\end{aligned}
$$

2. Benzene Emissions Estimation

EPA's NONROAD2004 model does not generate emissions estimates for benzene, or any other hazardous air pollutant. Benzene emissions were estimated outside of NONROAD by compiling engine-specific benzene speciation profiles from EPA's NMIM (EPA, 2005d). These speciation factors were then applied to the SCC-level VOC exhaust and evaporative VOC emissions output (after adjustment for the large S-I standard). Table VI-6 shows the benzene emission factors that were multiplied by the identified SCC-level VOC exhaust or evaporative emissions. Evaporative hydrocarbon emissions as calculated by NONROAD are comprised of crankcase, diurnal, spillage, and vapor displacement components. No benzene emission factors were available (or applied) for Liquified Petroleum Gas or Compressed Natural Gas-fired equipment.

B. EMISSION SUMMARIES

A Tier 3 summary of NONROAD model benzene emissions for each scenario is presented in Table VI-7. The final benzene emissions depend on the contribution of evaporative VOC relative to exhaust VOC emissions for a given SCC, since the benzene speciation profiles vary for these two components. Within the NONROAD model sector, gasoline or S-I engines are a much more important benzene source than diesel or compression-ignition engines. No benzene emission factors are available for LPG and CNG engines. For gasoline engines, lawn and garden, commercial, industrial, and recreational boats dominate. Construction and industrial engines are the top-emitting diesel equipment categories. All of these are emission categories that are affected by Federal standards established since the CAAA of 1990.

In the *without-CAAA* scenarios, overall benzene emissions increase between 1990 and 2000, and up through 2010 and 2020. This is due primarily to growth in equipment populations, though some categories show slight declines (e.g., gasoline construction and industrial equipment). For the *with-CAAA* scenarios, overall benzene emissions decrease between 1990 and 2000, decrease further in 2010, and level off somewhat in 2020. For example, VOC and benzene emissions increase somewhat between 2010 and 2020 for lawn and garden and light commercial gasoline engines, which leads to a slight overall increase. For a given time period, benzene emissions for specific nonroad categories either decrease or increase depending on the phase-in of Federal engine standards impacting VOC emissions, and the effects of category-specific growth rates.

Some of the uncertainties in this analysis include the accuracy of the benzene speciation profiles, which for some engines are based on highway vehicle testing, or limited nonroad engine test data. In addition, the NONROAD model runs used all default activity data for the counties in the Houston area (e.g., equipment populations, annual hours of use). The default activity may not be representative of actual nonroad activity for certain equipment types in these three counties, since NONROAD uses top-down allocation methods to estimate these variables.

Table VI-7. Houston Area NONROAD Model Benzene Emissions, tpy

TIER 2 NAME	TIER 3 NAME	1990	2000 *without-CAAA*	2000 *with-CAAA*	2010 *without-CAAA*	2010 *with CAAA*	2020 *without-CAAA*	2020 *with-CAAA*
Non-Road Gasoline	recreational	14.16	18.57	16.09	39.48	26.54	48.23	16.00
	construction	22.82	28.91	17.77	27.80	7.39	25.77	6.58
	industrial	77.26	62.69	38.12	53.69	17.44	44.78	14.63
	lawn & garden	411.45	510.78	304.86	645.70	180.38	787.09	202.09
	farm	0.26	0.27	0.19	0.28	0.12	0.29	0.09
	light commercial	122.35	163.09	90.46	228.45	58.56	296.90	73.43
	logging	0.02	0.03	0.02	0.04	0.02	0.06	0.02
	airport service	0.47	0.53	0.30	0.62	0.12	0.72	0.10
	railway maintenance	0.15	0.17	0.09	0.20	0.05	0.22	0.06
	recreational marine vessels	49.20	64.15	57.76	70.44	37.01	78.04	31.68
Non-Road Diesel	recreational	0.03	0.04	0.04	0.05	0.04	0.07	0.03
	construction	21.37	30.90	24.76	36.06	15.04	39.44	8.26
	industrial	12.76	10.83	8.50	11.74	4.16	13.21	2.48
	lawn & garden	1.59	2.17	1.99	3.31	1.58	4.51	1.21
	farm	0.97	0.80	0.73	0.72	0.41	0.73	0.23
	light commercial	4.02	5.08	4.71	6.99	4.09	8.96	2.73
	logging	0.01	0.01	0.01	0.01	0.00	0.01	0.00
	airport service	0.73	0.81	0.64	1.22	0.46	1.68	0.34
	railway maintenance	0.15	0.19	0.17	0.24	0.15	0.30	0.10
	recreational marine vessels	0.13	0.17	0.17	0.23	0.21	0.30	0.21
Other[1]	liquified petroleum gas							
	compressed natural gas							
	Total	**739.90**	**900.20**	**567.38**	**1,127.27**	**353.77**	**1,351.29**	**360.28**

NOTE: [1] No benzene emission factors available for LPG and CNG engines; emissions reported as null values.

CHAPTER VII. REFERENCES

Barrett, 2005: E. Barrett, Air Quality Planning and Implementation Division, Texas Commission on Environmental Quality, telecommunication with K. Thesing, E.H. Pechan & Associates, Inc., September 22, 2005.

BOC, 2004: Bureau of the Census, *American Housing Survey,* "Table 1A-4 Selected Equipment and Plumbing," U.S. Department of Commerce, Washington, DC, September 2004, available from http://www.census.gov/hhes/www/housing/ahs/nationaldata.html.

BOC, 2005: Bureau of the Census, "1997 NAICS and 1987 SIC Correspondence Tables," information accessed on July 21, 2005 from http://www.census.gov/epcd/www/naicstab.htm.

Broderick and Houck, 2005: Broderick, D.R. and J.E. Houck, OMNI Environmental Services, Inc., "Projected Residential Wood Burning Appliance Installations," July 18, 2005.

Chi, et al., 2004: Chi, R., A. Unal, D. Tian, and A. Russell, "NONROAD Emissions Model Uncertainty Analysis for the State of Georgia," Georgia Institute of Technology, presented at the International Emission Inventory Conference "Working for Clean Air in Clearwater," Clearwater, FL, June 7-10, 2004.

City of Houston, 2008: "Request for Correction of Information Under the Data Quality Act and EPA's Information Quality Guidelines," letter from Office of the Mayor, City of Houston, to the U.S. Environmental Protection Agency, Washington, DC, July 9, 2008.

DOE, 2003: U.S. Department of Energy, "Steam-Electric Plant Operation and Design Report," Form EIA-767, data files for 1985-2002, Energy Information Administration, Washington, DC, 2003.

EIA, 2005a: Energy Information Administration, "State Energy Consumption, Price, and Expenditure Estimates (SEDS)," U.S. Department of Energy, Washington, DC, Texas data retrieved on July 20, 2005 from http://www.eia.doe.gov/emeu/states/state.html?q_state_a=tx&q_state=TEXAS.

EIA, 2005b: Energy Information Administration, "Distribution and Production of Oil and Gas Wells by State," U.S. Department of Energy, Washington, DC, Texas data retrieved on July 22, 2005 from http://www.eia.doe.gov/pub/oil_gas/petrosystem/petrosysog.html.

EIA, 2005c: Energy Information Administration, "Sales of Distillate Fuel Oil by End Use," U.S. Department of Energy, Washington, DC, Texas data retrieved on July 20, 2005 from http://tonto.eia.doe.gov/dnav/pet/pet_cons_821dst_dcu_STX_a.htm.

EIA, 2005d: Energy Information Administration, *Annual Energy Outlook 2005 With Projections to 2025*, DOE/EIA-0383(2005), U.S. Department of Energy, Washington, DC, February 2005.

ENVIRON, 2004: ENVIRON International Corporation, “Final Report, CRC Project E-68, Analysis of EPA’s Draft Plan for Emissions Modeling in MOVES and MOVES GHG,” prepared for Coordinating Research Council, Inc., May 2004.

EPA, 1994a: U.S. Environmental Protection Agency, “Gasoline Distribution Industry (Stage I) – Background Information for Proposed Standards,” EPA-453/R-94-002a, Office of Air Quality Planning and Standards, Research Triangle Park, NC, January 1994.

EPA, 1994b: U.S. Environmental Protection Agency, “Gasoline Distribution Industry (Stage I) – Background Information for Promulgated Standards,” EPA-453/R-94-002b, Office of Air Quality Planning and Standards, Research Triangle Park, NC, November 1994.

EPA, 1999: U.S. Environmental Protection Agency, “Final Regulatory Impact Analysis: Control of Emissions from Marine Diesel Engines,” EPA420-R-99-026, Office of Mobile Sources, Ann Arbor, MI, November 1999.

EPA, 2002: U.S. Environmental Protection Agency, “Final Regulatory Support Document: Control of Emissions from Unregulated Nonroad Engines,” EPA420-R-02-022, Office of Mobile Sources, Ann Arbor, MI. September 2002.

EPA, 2004a: U.S. Environmental Protection Agency, Office of Transportation and Air Quality, NONROAD2004 Model, [Computer software], Ann Arbor, MI, May 2004. Available at: http://www.epa.gov/otaq/nonrdmdl.htm.

EPA, 2004b. U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, “Documentation For the Onroad National Emissions Inventory (NEI) For Base Years 1970-2002,” Research Triangle Park, NC, January 2004.

EPA, 2005a: U.S. Environmental Protection Agency, “Documentation for the 1990 National Emission Inventory for Hazardous Air Pollutants,” Emission Inventory Group, Research Triangle Park, NC, November 15, 2005.

EPA, 2005b: U.S. Environmental Protection Agency, “2002 EPA National Emission Inventory Point Source Database (Draft),” Office of Air Quality Planning and Standards, Emission Inventory Group, http://www.epa.gov/ttn/chief/net/2002inventory.html#point, February 2005.

EPA, 2005c: U.S. Environmental Protection Agency, Technology Transfer Network, http://www.epa.gov/ttn/chief/emch/projection/emshap30.html, Clearinghouse for Emission Inventories and Factors, Research Triangle Park, NC, accessed July 2005.

EPA, 2005d: U.S. Environmental Protection Agency, Office of Transportation and Air Quality, NMIM Software (NMIM20050311) and NMIM County Database (NCD20050427), Ann Arbor, MI, April 2005, available at: http://www.epa.gov/otaq/nmim.htm

EPA, 2006: U.S. Environmental Protection Agency, "WebFIRE" [web-based electronic emission factor database], Office of Air Quality Planning and Standards, Research Triangle Park, NC. Available from http://cfpub.epa.gov/oarweb/index.cfm?action=fire.main, accessed on January 3, 2006.

EPA, 2007: U.S. Environmental Protection Agency, "Control of Hazardous Air Pollutants from Mobile Sources – Regulatory Impact Analysis," EPA420-R-07-002, Assessment and Standards Division, Office of Transportation and Air Quality, February 2007.

ERG, 2005: Eastern Research Group, Inc., "Documentation for the Draft 2002 Point Source National Emissions Inventory," prepared for U.S. Environmental Protection Agency, Emission Inventory Group, Research Triangle Park, NC, February 7, 2005.

ERG and Pechan, 2004: Eastern Research Group and E.H. Pechan & Associates, Inc., "Documentation for the 2002 Electric Generating Unit (EGU) National Emissions Inventory (NEI)," prepared for U.S. Environmental Protection Agency, Emission Inventory Group, Research Triangle Park, NC, September 2004.

FAA, 2005: Federal Aviation Administration, *Terminal Area Forecasts (TAF)*, U.S. Department of Transportation, Washington, DC, retrieved on July 27, 2005 from http://www.apo.data.faa.gov/faatafall.HTM.

Frey and Bammi, 2002: Frey, H. and S. Bammi, "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, 52(4): 435-449, 2002.

Hanna, et al., 2004: Hanna, A., J. Vukovich, S. Arunachalam, D. Loughlin, H. Frey, J Touma, J. Irwin, and V. Isakov, "Assessment of Uncertainty in Benzene Concentration Estimates in the Houston, TX Area," *Proceedings, Annual Meeting of the Air & Waste Management Association*, June 22-25, 2004, Indianapolis, IN, published by Air & Waste Management Association, Pittsburgh, PA, June 2004.

Houyoux, 2004: Houyoux, Marc, U.S. Environmental Protection Agency, "CAIR Emissions Inventory Overview," Research Triangle Park, NC, July 23, 2004.

NARSTO, 2005: North American Research Strategy for Tropospheric Ozone, "Improving Emission Inventories for Effective Air Quality Management Across North America, A NARSTO Assessment," prepared by the NARSTO Emission Inventory Assessment Team, 2005.

NRC, 2000: National Research Council, "Modeling Mobile-Source Emissions," National Academy Press, Washington, DC, 2000.

Pechan, 2005a: E.H. Pechan & Associates, Inc., "Emission Projections for the Clean Air Act Second Section 812 Prospective Analysis – Draft Report," prepared for Industrial Economics, Inc., November 2005.

Pechan, 2005b: E.H. Pechan & Associates, Inc., "Crosswalk Between Maximum Achievable Control Technology (MACT) Codes and Emission Activity Growth Factors, Final Technical Memorandum," prepared for U.S. Environmental Protection Agency, April 15, 2005.

Pechan, 2005c: E.H. Pechan & Associates, Inc., "Crosswalk Between Source Classification Code (SCC)s and Emission Activity Growth Factors, Final Technical Memorandum," prepared for U.S. Environmental Protection Agency, September 9, 2004.

Pechan, 2006: E.H. Pechan & Associates, Inc., "Documentation for the Final 2002 Nonpoint Source National Emission Inventory for Criteria and Hazardous Air Pollutants," prepared for U.S. Environmental Protection Agency, Emission Inventory Group, Emissions, Monitoring, and Analysis Division, Office of Air Quality Planning and Standards, Research Triangle Park, NC, forthcoming (expected March 2009).

Price, 2005: Price, Jason, Industrial Economics, Inc., *BenMAP_812_target_year_pop_projections.xls* [Electronic File], Cambridge, MA, transmitted to Andy Bollman, E.H. Pechan & Associates, Inc., July 7, 2005.

Radian, 1995: Radian Corporation, "Development Of The Houston Area Source Toxic Emissions (HASTE) Inventory, Final Report," prepared for Texas Natural Resource Conservation Commission Austin, Texas. February 1995.

Rubick, 2005: C. Rubick, Texas Commission on Environmental Quality, Excel spreadsheet entitled, *Revised VOC Distribution And Benzene Emissions From Marine Loading*, provided to K. Thesing, E.H. Pechan & Associates, Inc., September 21, 2005.

Sierra, 2004: Sierra Research, Inc., "Evaluation of MOBILE Models: MOBILE6.1 (PM), MOBILE6.2 (Toxics), and MOBILE6/CNG," prepared for National Cooperative Highway Research Program, 2004.

Starcrest, 2000: "Houston-Galveston Area Vessel Emissions Inventory," Prepared for Port of Houston Authority, Texas Natural Resource Conservation Commission, prepared by Starcrest Consulting Group, LLC, Houston, TX, November 2000.

TRB, 2007: Transportation Research Board, "Metropolitan Travel Forecasting – Current Practice and Future Direction," Special Report 288, Washington, DC, 2007.

Truex and Norbeck, 1998: Truex, Dr. T.J. and Dr. J.M. Norbeck "Evaluation of Factors That Affect Diesel Exhaust Toxicity," University of California-Riverside, Center for Environmental Research and Technology. Riverside, CA. 1998.

Woods & Poole, 2001: Woods & Poole Economics Inc., "Population by Single Year of Age," CD-ROM. 2001.

APPENDIX B:

BENZENE AIR QUALITY AND EXPOSURE MODELING REPORT

APPENDIX B

Air Quality and Exposure Modeling in Support of Section 812 Benzene Case Study: Methodology and Results

Table of Contents

Table of Contents

List of Tables

List of Figures

List of Figures

List of Acronyms

AERMAP	American Meteorological Society/EPA Regulatory Model Terrain Preprocessor
AERMET	AERMOD meteorological preprocessor
AERMOD	American Meteorological Society/EPA Regulatory Model
ASPEN	Assessment System for Population Exposure Nationwide
CAA	Inventories with Clean Air Act controls included
CALPUFF	Lagrangian puff model
CAS	Chemical Abstract Service
CHAD	Consolidated Human Activity Database
DEM	Digital Elevation Model
EMS-HAP	Emissions Modeling System for Hazardous Air Pollutants
EPA	United States Environmental Protection Agency
FAA	Federal Aviation Administration
FEMA	Federal Emergency Management Association
FIPS	Federal Information Processing Standards
FSL	Forecast Systems Laboratory
HAP	Hazardous Air Pollutant
HAPEM5	Hazardous Air Pollutant Exposure Model, Version 5
HAPEM6	Hazardous Air Pollutant Exposure Model, Version 6
ISCST3	Industrial Source Complex Short Term Model
ISHD	Integrated Surface Hourly Database
MACT	Maximum Available Control Technology standards for HAP, established under Section 112 of the Clean Air Act
MOBILE6	Mobile source emissions model, Version 6
NATA	National Air Toxics Assessment
NEI	EPA's National Emission Inventory
NON-CAA	Inventories without Clean Air Act controls included
NONROAD	Nonroad source emissions model
NWS	National Weather Service
OAQPS	EPA's Office of Air Quality Planning and Standards
OTAQ	EPA's Office of Transportation and Air Quality
SAF	Spatial Allocation Factor
SAMSON	Solar and Meteorological Surface Observation Network
SAROAD	Storage and Retrieval of Aerometric Data: Air pollution chemical species classification
SCC	Source Classification Code
SCRAM	Support Center for Regulatory Atmospheric Modeling
SIC	Standard Industrial Classification code used for Federal economic statistics
TAF	Terminal Area Forecast
USGS	United States Geological Survey
UTM	Universal Transverse Mercator
VMT	Vehicle Miles Travelled
WRPLOT	Wind Rose PLOT

1 Introduction

The U.S. Environmental Protection Agency's Office of Air and Radiation is working on a case study to estimate the human health benefits of reducing benzene emissions in the Houston, Texas, area. This document details the emissions processing, air quality modeling, and exposure modeling in support of this case study. A total of seven scenarios were processed and modeled. Listed below are the inventories and how they will be referenced throughout this report:

1. Year 1990: 1990
2. Year 2000 with Clean Air Act controls: 2000 CAA
3. Year 2000 without Clean Air Act controls: 2000 NON-CAA
4. Year 2010 with Clean Air Act controls: 2010 CAA
5. Year 2010 without Clean Air Act controls: 2010 NON-CAA
6. Year 2020 with Clean Air Act controls: 2020 CAA
7. Year 2020 without Clean Air Act controls: 2020 NON-CAA

The purpose of the modeling was to show how the Clean Air Act Amendments of 1990 have affected historical, and will affect projected, ambient benzene concentrations. EPA accomplished this task by modeling the air quality impact of emissions inventories with and without emission controls required by the Clean Air Act Amendments. The methodology used in this analysis is consistent with that applied in previous EPA studies (U.S. EPA, 2002a). The emissions and air quality modeling described in this report were based upon emissions projections developed by Pechan (2006). (See Appendix A.)

The emphasis of the study was the Houston metropolitan area, specifically Brazoria, Galveston and Harris counties. The domain of the modeling study is shown in Figure 1. The three counties of emphasis are shown in red outlines and the Houston metropolitan area is shown in yellow. Key roads are also shown.

In this study, for each of the seven modeling cases above, the total emissions inventory was contained in four separate inventories: 1) point, 2) nonpoint, 3) onroad, and 4) nonroad. The point inventory contained both "major" and "area & other" sources that had reasonably known location coordinates. The nonpoint inventory contained county-level area & other emissions by Source Classification Codes (SCC). These emissions were spatially allocated to census tracts or airports during emissions processing to further refine their location. Emissions related to airports such as aviation gasoline distribution were allocated to the airports because the locations of the airports were known. Other nonpoint emissions were allocated to the census tracts. The onroad inventory contained link-based emissions where the links represent road segments. Using link-based emissions allowed the user to have more accurate locations of the onroad emissions as opposed to using spatial allocation to census tracts. The nonroad inventory contained nonroad mobile emissions at the county and SCC level that were spatially allocated to the census tracts or airports (e.g., aircraft and airport support equipment) in similar fashion as the nonpoint emissions.

Once emissions processing was completed, the emissions were processed through the American Meteorological Society/EPA Regulatory Model (AERMOD) to get annual average concentrations. These concentrations were then processed through the Hazardous Air Pollutant Exposure Model, Version 6 (HAPEM6) to yield exposure concentrations. Figure 2 shows the steps involved from initial emissions processing to HAPEM6 exposure concentrations.

The report is divided as follows:

Section 2: Modeling methodology (model selection, averaging period, treatment of terrain, etc.);
Section 3: Inventory preparation for EMS-HAP;
Section 4: Processing in EMS-HAP and emissions summaries;
Section 5: AERMOD results;
Section 6: HAPEM6 results;
Section 7: Limitations; and
Section 8: Conclusions.

2 Modeling Methodology

This section presents the methodology used for the air quality modeling in this study and follows the general methodology discussed in previous U.S. EPA analyses of urban areas (U.S. EPA, 2002a).

2.1 Model Selection

The air quality model chosen for this study was the American Meteorological Society/U.S. EPA Regulatory Model (AERMOD) dispersion model (U.S. EPA 2004a)[1]. AERMOD handles multiple sources, incorporates building downwash, has flexibility in receptor location choices, models the effects of complex terrain, and models the effects of wet or dry deposition. AERMOD also includes the option to vary emissions by season and hour of day.

2.2 Averaging Period

AERMOD calculates an hourly concentration at each receptor and other averaging periods can also be calculated. The averaging period selected is based on the intended use of the model output. To estimate chronic exposure to benzene, annual average air concentrations are generally needed for this study. In addition, we calculated hourly average concentration for later calculation and input into HAPEM6.

2.3 Receptor Strategy

A receptor is any location where ambient concentration estimates are needed. AERMOD utilizes user supplied receptor locations. For this study, receptors were chosen as the census block group centroids[2]. For the 1990 simulation, the receptors were the 1990 census block group centroids (Figure 3), giving a total of 2,429 receptors. For all other AERMOD simulations in the study (2000, 2010, and 2020), the 2000 census block group centroids were chosen as the receptors (Figure 4), for a total of 2,285 receptors. Table 1 lists the year of choice for the receptors for each AERMOD simulation. On the whole, for both years, a large number of the receptors were concentrated around the Houston metropolitan area instead of the outlying areas. Overall, the spatial distribution of receptors appeared similar for both 1990 and 2000.

2.4 Terrain

Terrain elevation at each source and receptor are needed input into AERMOD. Source and receptor elevations are user input or are determined by American Meteorological Society/EPA Regulatory Model Terrain Preprocessor (AERMAP) using U.S. Geological Survey Digital

---

[1] In Sections 3 and 4, emissions processing refers to processing for the Industrial Source Complex Short Term Model (ISCST3). This is because EMS-HAP is currently configured to create ISCST3 emissions files. Since ISCST3 and AERMOD emissions files are formatted similarly, EMS-HAP can be used to create AERMOD emissions files. In Sections 3 and 4, when ISCST3 is used, it is understood to be for AERMOD as well.

[2] Census block groups are geographic entities within the same census tract. Block groups usually contain between 250 and 550 housing units.

Elevation Model (USGS DEM) terrain data. For the sources, if USGS data are not available, EMS-HAP can assign elevation to sources based on census tract centroid elevations (U.S. EPA, 2004b). Since the terrain is relatively flat over the Houston area (Figure 5), the flat terrain option was chosen for the AERMOD simulations.

2.5 Land use

Land use data were used to designate sources as urban or rural for dispersion modeling purposes. For 1990 tracts (used in the 1999 AERMOD run), a tract was considered urban if the tract population density from the 1990 census was greater than 750 people per km^2 (U.S. EPA, 2002b). For year 2000 tracts (to be used for all other years), urban and rural designations were based on a combination of population and commercial/industrial land use. A tract was considered urban if: 1) the residential population density based on the 2000 census was greater than 750 people per km^2, based on using the tract land area, not total tract area (land area plus water area); or 2) The area of buildings classified as commercial, industrial, or institutional based on data from the Federal Emergency Management Agency (FEMA) exceeds 50% of the total tract area (land plus water) (U.S. EPA, 2004b). For the 1990 AERMOD simulation, the 1990 census tract designations were used while the 2000 census tract designations were used for the other simulations. The 1990 designations are shown in Figure 6 and the 2000 designations are shown in Figure 7. It can be seen that that the distribution of urban and rural tracts was similar between 1990 and 2000 with the majority of urban tracts near Houston.

When processing emissions, we used urban/rural classification of tracts to assign emissions estimates. The urban/rural designation is important for AERMOD modeling when assigning deposition parameters. For the nonpoint and nonroad sources, excluding airport emissions, the sources were assigned the urban or rural designation of the census tracts to which the emissions were assigned during spatial allocation in EMS-HAP. For point sources, each point source was assigned the urban/rural designation of the closest tract. For point, nonpoint, and nonroad sources (excluding airport related emissions), the 1990 census tract designations were used for the 1990 AERMOD simulation, and the 2000 census tract designations were used for the other simulations (Table 1). For all simulations, the link based onroad emissions were modeled as rural sources. This is consistent with previous studies in Houston (U.S. EPA, 2002a). For nonpoint and nonroad airport related emissions, the sources were modeled as rural sources.

If urban sources are being processed, AERMOD requires an urban population to be input into the model as well (U.S. EPA, 2004a) for modeling urban heat island effect on dispersion. For the urban population, the Harris County population was used since this represents the bulk of the Houston metropolitan area. For the 1990 AERMOD simulation, the 1990 census population of 2,800,000 was used. For all other simulations, the 2000 census population of 3,400,000 was used. Population estimates can be found at the U.S. Census website, www.census.gov.

2.6 Meteorological Data

Meteorological data were prepared for two years, 1990 and 2000. The 1990 meteorological data are used for the 1990 AERMOD simulation. The year 2000 data are used for the other simulations. The AERMOD model requires values of meteorological variables shown in Table 2. The data are derived from a combination of a selected surface station and an upper air station. Also included with the data in Table 2 are the year, month, day, hour of day, and Julian day, all in local standard time (LST). For all AERMOD runs, the surface roughness length was set to 0.6 meter.

2.6.1 Selection of surface and upper air stations

For 1990, the only surface station available was George Bush Intercontinental Airport (IAH), located north of Houston; therefore, it was chosen as the representative station for 1990. For 2000, there were two stations to consider, IAH and Hobby Field (HOU) located in southern Houston (Fig. 8a). For 2000, HOU reported fewer calms (858 hours) and missing data than IAH (1,554 hours). Comparing annual average values of temperature, wind speed, wind direction, and rainfall, the two stations showed very little difference. In AERMOD, hours with missing meteorological data or calm winds are given concentrations of zero. When AERMOD calculates the annual average concentration, those hours are not included in the averaging. This could lead to an underestimate of the annual average concentration for the station with more missing hours when comparing results using the two stations. Given the fewer number of calms and missing hours for HOU in 2000, it was decided to use HOU as the surface station for year 2000.

The upper air station selected was based on the station considered the most representative of the city. Two stations were considered: Victoria and Lake Charles. However the Victoria station moved to Corpus Christi in 1990, which was nearer the Gulf Coast than Houston; therefore, Lake Charles was chosen as the representative stations for both 1990 and 2000. The relative location and distance between the surface stations and Lake Charles upper air station are shown in Figure 8b. Table 1 summarizes the years used for each AERMOD simulation.

2.6.2 Meteorological data pre-processing

The AERMOD Meteorological Preprocessor (AERMET) (U.S. EPA, 2004c) was used to process the National Weather Service (NWS) data for both 1990 and 2000. For 1990, the surface data was in the Solar and Meteorological Surface Observation Network (SAMSON) format while for 2000 the surface data was in the Integrated Surface Hourly Database (ISHD) format. For both years, the Lake Charles upper air data was in the Forecast Systems Laboratory (FSL) format. The AERMET preprocessor reads surface and upper air data and creates files of meteorological variables needed by AERMOD. AERMET has three stages. The first reads the surface and upper air data files from the user and performs several quality assurance checks of the data for missing values and values considered out of range by the user. The second stage merges the surface and upper air data into one dataset while the third stage reads the merged data, calculates necessary boundary layer variables and creates the surface and profile files used in AERMOD. For a detailed discussion of the stages and boundary layer calculations, see the AERMET User's Guide (U.S. EPA, 2004c).

2.6.3 Climatological comparisons

Figure 9 shows wind roses for an eight-year period (1984-1992) for IAH, IAH in 1990 and HOU in 2000. An eight-year climatological period was chosen because the data for these data are readily available from EPA's Support Center for Regulatory Atmospheric Modeling (SCRAM) (www.epa.gov/scram001) and are easily imported in the graphical program Wind Rose PLOT (WRPLOT) (also available from SCRAM), which is used to generate the wind roses. For 1990 and 2000, the winds appeared to be similar with the climatological record in that the winds are generally from the southeast. HOU for 2000 had a slightly higher frequency of southerly winds than the climatology for IAH in 1990. Table 3 shows the annual average wind speeds and directions as well as annual rainfall, average minimum and maximum daily temperature for the two years and the climatological record. Again, the two years' data were similar to the climatological record. Note that in Table 3, the values shown for 1990 and 2000 are for the post AERMET data, since these data were input into AERMOD.

2.7 Background

Background concentrations used in the study for all years and modeling scenarios were the 1999 county specific background concentrations as used for the 1999 NATA (Table 4). For details about the 1999 background values see http://www.epa.gov/ttn/atw/nata1999/background.html or Battelle (2003). Background concentrations are added to AERMOD modeled concentrations at each receptor (block group centroids) in post-processing and input into HAPEM6 exposure modeling. Every receptor in a particular county receives the same background concentration.

2.8 Model Evaluation

Model evaluation is performed by comparing modeled concentrations to observed concentrations. In addition to the census block group centroids, daily and annual average model concentrations were calculated at monitor locations. In 1990, there was only one benzene monitor available in the domain, which is not enough for a valid comparison. In 2000, there were 15 monitor locations available for comparison, mostly in southern Harris County (Figure 10). Model-to-monitor comparisons were only performed for the year 2000 AERMOD results. Monitor observations were obtained from EPA's Air Toxics Archive, which contains multiple years of monitor observations for multiple HAPs and across the U.S. The Archive contains a program that performs quality assurance on daily monitor observations and calculates an annual average concentration for each valid monitor.

3 Emissions Pre-processing: Inventory Modification for EMS-HAP

The stationary source inventories for the case study were developed from the 1990 National Emissions Inventory (NEI) and the 2002 draft NEI. The onroad mobile inventories were developed from the Mobile Source Emissions Model, version 6 (MOBILE6) (Cook and Glover, 2002) and the nonroad inventories (excluding aircraft, locomotives, and commercial marine vessels) were developed from the Nonroad Source Emissions Model (NONROAD) model (U.S. EPA, 2004d). The development (application of growth factors and controls, etc.) of the inventories is detailed in Pechan (2006).

We developed seven inventory scenarios, as described in Section 1 (Introduction). Below we detail the inventory development undertaken for each inventory category. The inventories developed by Pechan (2006) required some modifications before input into Emissions Modeling System for Hazardhous Air Pollutants (EMS-HAP). These modifications included the development of some source characteristics needed by AERMOD and are described further below.

3.1 Point Sources

For the point inventories, the required variables for input into EMS-HAP are shown in Tables 3-6 and 3-7 in Chapter 3 of the EMS-HAP User's Guide (U.S. EPA, 2004b). According to the EMS-HAP User's Guide, records in the inventory should be unique by State and county Federal Information Processing Standards (FIPS), a site identifier for the source, the emissions release point identifier for the stack at the site, and a unique pollutant identifier (CAS). This means each stack at a facility is a unique record. Initially, this was not the case for the point inventories, becuase many of the site identifiers were blank. Therefore, the site identifier variable for all sites was changed based on two cases. For the first case, if the site identifier was blank or missing, the site identifier was set equal to the concatenation of the state and county FIPS code, a dash, and the record number. For example, if the site was in the third record in the inventory and the FIPS code was 48039, the site identifier became 48039-0003. If the site identifier was not blank, the site identifier was the concatenation of the state and county FIPS code, a dash, and the last 15 characters of the site identifier. The emissions release point identifier was just the concatenation of the emission release point identifier already in the inventory and the record number of the source. The changes to the site identifier and the emissions release point identifier ensured each record was unique.

As can be seen in Table 3-6 of the EMS-HAP User's Guide, a variable called SRC_TYPE is needed to distinguish the site as a major source or area & other source in order to group emissions into categories for later use in AERMOD. Initially, the source type was missing in the inventories. Normally this variable can be obtained from the facility category variable in the NEI inventories. In order to assign source types to the point sources, the emissions inventories were merged with a reference inventory that contained the source types. The merger was done by assigning the source type of the closest source from the reference inventory to the Houston inventories. Hypothetically, the merged sources would represent the same site. Indeed, after the merger it was found that, for each Houston source, the reference source was the same site and distances between the sites was always less than 1 meter. For the 1990 and NON-CAA inventories for 2000, 2010, and 2020, the reference inventory was the 1990 inventory. This inventory was chosen because the 1990 and NON-CAA inventories were based on the 1990 inventory (Pechan, 2006). For the CAA inventories for 2000, 2010, and 2020, the reference inventory was the draft 2002 NEI inventory, because this was the basis of the Clean Air Act inventories (Pechan, 2006).

Once the source types were added, volume source characteristics were assigned to sources that would best be modeled as volume sources, mainly fugitive emissions, as recommended by the EMS-HAP User's Guide (U.S. EPA, 2004b). Sources chosen to be modeled as volume sources were based on the SCC code of the sources. These SCC codes are shown in Table 5. These sources were modeled with horizontal and vertical dispersion parameters, $\sigma_x = 1.5$ meters, $\sigma_z = 3$ meters, and a release height of 2 meters.

Next, several sources in the 1990 and NON-CAA inventories needed location adjustments because of coordinate issues. The coordinates of the sources listed in Table 6 were adjusted using average coordinates of other stacks associated with the same site (i.e., same site ID). The other stacks were generally clustered around each other with the sources listed in Table 6 located away from the other stacks. Once the adjustments were done on the locations, all non-zero emissions were output to SAS datasets, ready for processing in EMS-HAP.

3.2 Nonpoint and Nonroad Emissions

The nonpoint and nonroad inventories were converted from Microsoft Access to SAS with only non-zero or non-missing emissions retained (to save computer space). The inventories contained the necessary variables for input into EMS-HAP as shown in Table 2-3 for the nonpoint inventory and Table 2-4 for the nonroad inventory in the EMS-HAP User's Guide (U.S. EPA, 2004b). The only changes done to the inventories were to move aircraft, locomotive, and commercial marine vessel emissions from the nonpoint inventories to the corresponding nonroad (year and Clean Air Act scenario) inventories. This was done so that all nonroad emissions for a given year and Clean Air Act scenario would be together and separate from the nonpoint emissions to avoid any confusion later. The aircraft, locomotive, and commercial marine vessel emissions were initially separate from the other nonroad emissions, because they were processed differently (Pechan, 2006) from other nonroad sources. Once these changes were made, the inventories were output to SAS datasets, ready for EMS-HAP processing.

3.3 Onroad

Each onroad inventory for each modeling scenario contained seasonal-hourly link-based emissions, allowing them to be processed through EMS-HAP as AERMOD area sources using two of the point programs (PtTemporal and PtFinal_ISCST3). Accordingly, the required variables for input into EMS-HAP are shown in Tables 5-2 and 5-3 in Chapter 5 of the EMS-HAP User's Guide (U.S. EPA, 2004b).

Several issues were found during the course of inventory quality assurance (QA) and preparation for input into EMS-HAP. The inventories contained emissions for some counties not included in the overall study and were removed before further QA and processing. Also, some seasonal-hourly emissions values were missing because these time/link instances had no vehicle miles traveled (VMT) provided from the Texas Transportation Institute (TTI) and could be assumed to be zero, according to email communication from Pechan (2006).

Several necessary variables in the provided inventories had no initial values, requiring values to be defaulted or created. These are the same variables as listed in Tables 3-6 and 3-7 of the EMS-HAP User's Guide (U.S. EPA, 2004b) for point sources (SAROAD, NTI_HAP, SRC_TYPE, ISCTYPE, EMRELPTY, AINPLUM, AND ARELHGT). Somewhat similar to the point sources, the site identifier (SITE_ID), which is supposed to be unique for each individual link (also called a road segment) was always blank in the onroad inventories. Initially, the only way to distinguish road segments was by the Universe Transverse Mercator (UTM) coordinates (the "southwest" corner of the ISCST3 area source). Refer to Figure 2-5 of the EMS-HAP User's Guide (U.S. EPA, 2004b) for a graphical description of the angular relationships for ISCST3 area sources. All road segments were given a temporary SITE_ID by simply concatenating the UTMX and UTMY variables with an underscore. The UTM coordinates were found to be in the incorrect UTM zone for Houston and had to be converted from Zone 14 to Zone 15 for consistency with the other source categories, after which new unique, intermediate SITE_IDs were created by concatenating the five-digit FIPS code with an arbitrary number.

According to the EMS-HAP User's Guide (U.S. EPA, 2004b), the aspect ratio (i.e., road length to road width, or AXLEN to AYLEN) of an AERMOD area source should not exceed 100:1. Each road segment with an excessive aspect ratio was split into the appropriate number of segments to satisfy the 100:1 limit. Each new segment was given a final SITE_ID, consisting of the intermediate SITE_ID concatenated with the ordered number of the split segments (i.e., SITE_ID-1, 2, 3…). The EMRELPID was set equal to the final SITE_ID concatenated with '__scc-' and the original 10-digit SCC code.

Some road segments had emissions given for both directions of traffic flow (Pechan 2006), which were then summed by SITE_ID and SCC. The road widths (AYLEN) for these segments had to be doubled to compensate for the subsequent consolidation of emissions records. Some road widths (AYLEN) with initial values of zero were given default values of 7.3152 meters (approximately the equivalent of a two-lane road, where each lane is 12 feet or 3.6576 meters across). According to Pechan (2006), zero value AYLEN with nonzero AXLEN was how TTI coded the local or intrazonal VMT. All other road segments had nonzero initial values of AYLEN, which does not appear to be a unique function of road type from the original 10-digit SCC code.

4 EMS-HAP Processing

This section details the processing of emissions through EMS-HAP. A brief explanation of EMS-HAP is given in Section 4.1. Modification or development of ancillary files needed by EMS-HAP follows in Section 4.2 and the actual processing by EMS-HAP is described in Section 4.3. Some post-processing of EMS-HAP output needed for final input into AERMOD is described in Section 4.4. Emissions summaries and maps are in Section 4.5.

4.1 EMS-HAP

The Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP Version 3.0) is a series of SAS based programs that process emissions inventories for input into subsequent air quality modeling. Currently EMS-HAP can create emissions input files for the Assessment System for Population Exposure Nationwide (ASPEN) dispersion model (U.S. EPA, 2000) or the Industrial Source Complex Short Term Version 3 (ISCST3) (U.S. EPA, 1995) dispersion model. With some minor post-processing by the user, ISCST3 ready files can be made ready for AERMOD.

During emissions processing EMS-HAP performs the following functions (U.S. EPA, 2004b):

- checks inventory location data, converts to latitude/longitude coordinates and removes inventory records with missing or out-of-range location data;
- checks inventory stack parameter data and defaults missing or out-of-range data;
- identifies three AERMOD point source types: point, volume, and area;
- groups and/or partitions individual pollutant species (e.g., groups lead oxide, lead nitrate into a lead group; partitions lead chromate into lead and chromium groups);
- where desired, further speciates individual pollutants (e.g., chromium and compounds into hexavalent chromium) by inventory MACT, SIC, or SCC code;
- facilitates the selection of pollutants and pollutant groups for modeling;
- assigns building heights and widths to certain stacks;
- spatially allocates county-level stationary and mobile source emissions to the census tract level or by grid cells;
- allocates certain county-level sources to particular locations (e.g., airports) to be modeled as AERMOD area sources with specific southwest corner, horizontal and vertical dimensions and angle;
- temporally allocates annual emissions to seasonal and day-type specific hourly emission rates to account for diurnal, day-of-week and seasonal patterns in emissions and imparts a day-type variation to MOBILE6.2-based seasonal and hourly emissions when processing for AERMOD;
- produces emission files formatted for direct input into the ASPEN model or, when processing for AERMOD, produces the Source (SO) pathway (emission-related inputs) of an AERMOD run stream.

More details about individual EMS-HAP programs follow in Section 4.3.

4.2 Modification of Ancillary Files

Following are details of modification of ancillary files needed by EMS-HAP. An ancillary file is a data file, text or binary, used by EMS-HAP to process emissions. Ancillary files contain data such as temporal allocation information (to allocate annual emissions to hourly or seasonal emissions) and spatial allocation factors (to allocate county level emissions to census tracts). Other ancillary files may include data about counties, such as zip codes, or information about pollutants.

4.2.1 Modification of temporal and surrogate ancillary files

Each inventory was checked to see if the temporal allocation files (which allocate annual emissions to season, day of week, and hour based on the SCC), the surrogate cross reference file (which assigns a spatial surrogate to emissions by SCC), and the source group file (which bins the emissions into groups by SCC) needed to be updated (i.e., adding SCC not already in the files). During emissions pre-processing, the point, nonpoint, and nonroad inventories were checked to see if any SCC needed to be added to the temporal allocation file. SCC codes that were added are shown in Attachment B-1, Table B-1. The nonroad and nonpoint inventories were checked to see if any SCC were to be added to the surrogate cross reference files. Two surrogate cross reference files were used, one for 1990 and one for all other years. Different files were used because the 1990 nonpoint and nonroad inventories used the 1990 census tract level surrogates (described in U.S. EPA, 2002b) while the other years (2000, 2010, and 2020) used the year 2000 census tract level surrogates (described in U.S. EPA 2004b). The SCC added to the 1990 surrogate cross reference file are shown in Table B-2 in Attachment B while SCC for the other years are shown in Table B-3. SCC often used surrogates of related SCC codes already in the surrogate cross-reference files.

For the onroad inventories, it has been previously found that many of the onroad SCC fit along a few temporal profile curves. Instead of listing all of the onroad SCC in the temporal profile, the onroad inventories are assigned to four new SCC based on fuel (gas, diesel) and road (interstate, local) types, for simplicity of the temporal profiles. Over 40% of the SCC codes found in the onroad inventories did not have matching temporal profiles in the EMS-HAP temporal allocation file. Previous work has suggested that for mobile inventories already temporally allocated by season and hour, onroad temporal profiles could be summarized by five basic profiles (U.S. EPA, 2004b). The profiles are shown in Attachment B, Figure B-1. Only two of those (non-diesel PM local/collector/arterial and non-diesel PM interstate/other freeway) are applicable for this benzene case study. Four new SCC were created for subsequent processing through EMS-HAP: 1) GAS_INTRST (gasoline engines; interstate roads), 2) GAS_LOCAL (gasoline engine; local roads), 3) DIE_INTRST (diesel engines; interstate roads) and 4) DIE_LOCAL (diesel engines; local roads). The two “LOCAL” SCC have the same temporal allocation factors, while the two “INTRST” SCC codes share a different set of TAFs (temporal profile). Gasoline and diesel SCC were divided here for any later analysis of fuel type contributions. The road segment emission records were then summed by final SITE_ID and new SCC. The assignment of SCC to the new temporal profile SCC is shown in Attachment B in Tables B4-B7.

4.2.2 Development of airport-related ancillary files

Nonpoint emissions such as aviation gasoline distribution and nonroad emissions such as aircraft and airport support equipment can be modeled from the airport locations rather than allocated to census tracts as other nonpoint and nonroad sources. This is because the location and sizes of airports are well known. In the Houston domain, there were nine airports, listed in Table 7 and shown in Figure 11.

The allocation of airport emissions takes place in the EMS-HAP module COPAX. (See Chapter 2, EMS-HAP User's Guide (U.S. EPA, 2004b).) In COPAX, the airport-related emissions are extracted from the nonpoint or nonroad inventory and allocated to airport locations using what are called airport allocation factors. This is discussed in Section 4.3.2.

4.2.2.1 Development of airport allocation factors

In order to allocate the airport emissions, airport allocation factors were developed for each of the study years for these airports. Airport allocation factors are used to allocate county-level emissions that are associated with airports (i.e., aircraft, airport support equipment, aviation gasoline distribution) and allocate them to airports in the county. Four types of allocation factors are used: general aviation, commercial aircraft, air taxi, and military aircraft. These factors are applied to specific SCC in the inventories. Table 7 lists the SCC and allocation factor types used for each SCC.

As stated, airport allocation factors were developed for each of the study years. The allocation factors are based on the Federal Aviation Administration's Terminal Area Forecast (FAA's TAF) model data (FAA, 2004), http://www.apo.data.faa.gov/, which can be run for different years. The TAF model calculates landing and take-off data for several aircraft types: general aviation, commercial aircraft, air taxi, and military aircraft at each airport for a given year. To calculate the allocation factors for EMS-HAP, the landing and take-off data or itinerant data were summed across all the airports for a given county for each of the allocation types. Next, each airport's individual allocation was calculated by dividing the airport's itinerant data for the allocation type (general aviation, commercial, etc.) by the county total for that allocation type. The result is a ratio that can be applied to a county-level airport emissions record to allocate the emissions to individual airports within the county:

$$Factor_{I,X} = \frac{IT_{X,I}}{IT_{X,county}} \qquad \text{(Eq. 1)}$$

where $\text{Factor}_{\text{I, X}}$ is the allocation factor for airport I, for X type (general aviation, commercial aircraft, air taxi, or military aircraft), $\text{IT}_{\text{X, I}}$ is the itinerant data for airport I for type X, and $\text{IT}_{\text{X, county}}$ is the county total of the itinerant data X. If the county sum for a particular type was zero then another allocation factor's itinerant data were used. For example, if the general aviation itinerant data were zero for a county, then the commercial aviation data were used. Other variations were used for other allocation types. For a complete description of the development of airport allocation factors see the EMS-HAP User's Guide (U.S. EPA, 2004b). Table 8 lists the airports in the Houston domain along with the allocation factors for each aviation type.

4.2.2.2 Airport areas

In an air toxics study, airport emissions are often run as AERMOD area sources[3], usually as a rectangular area, in AERMOD. For emissions processing in EMS-HAP, dimensions of the airport, as well as an orientation angle, are needed to create the area source. See the AERMOD User's Guide for details about necessary parameters of area source releases. In order to accurately represent airport emissions, as much of the runway and airport terminals should be covered by the rectangle. To determine the dimensions of the airports for the study, aerial photographs from ArcGIS of the airports were used and overlaid with rectangles, taking care to cover the terminals. Also, the orientation angle of the airports relative to north was calculated.

4.2.3 Source group assignments

In EMS-HAP, emissions can be assigned or binned into source groups, which are useful for assessing the emissions and air quality eminating from broad classes of sources. Emissions can be assigned by source sector (i.e., major, area & other, onroad, or nonroad) or emissions of related MACT, SIC, or SCC codes can be grouped together (e.g., onroad gasoline emissions, gasoline stations). For this study emissions were grouped into five source groups: major (0), area & other (excluding gasoline stations) (1), onroad mobile (2), nonroad mobile (3), and gasoline stations (4). The numbers in parentheses represent the group number assigned to the source groups.

In the point inventory, emissions were assigned to the major or area & other groups by the SRC_TYPE variable in the inventory. If SRC_TYPE equaled 'MAJOR', the emissions were assigned to group 0; otherwise the emissions were assigned to group 1. Similarly in the nonpoint airport emissions inventory and nonroad airport emissions inventory, both of which are treated as point inventories by EMS-HAP are assigned to groups by the SRC_TYPE variable, which is assigned during the module COPAX (U.S. EPA, 2004b).

In the remaining nonpoint and nonroad emissions inventories, emissions were assigned to groups based on the SCC of the emissions record. An ancillary file of EMS-HAP contains a list of the SCC from the nonpoint and nonroad inventories with an assigned group number. For the nonroad inventory, all emissions were assigned to group 3. For the nonpoint inventory, emissions were assigned to group 1 if not a gasoline station SCC, or group 4 if a gasoline station SCC. Table 9 lists the gasoline station SCC in the nonpoint inventories.

4.2.4 Tract polygon coordinates

A new feature added to EMS-HAP, Version 3 is to allow the user to create ISCST3 or AERMOD polygon area sources (keyword AREAPOLY in U. S. EPA, 2004a). In EMS-HAP Version 2,

---

[3] In this case, “area” does not refer to the emissions source group “area & other,” but as the type of emissions source as used by AERMOD (e.g., point, volume, area).

when processing emissions for AERMOD, the user was required to have a grid developed to which the county-level nonpoint or nonroad emissions would be allocated. Normally such a grid is 1 x 1 km and requires processing in programs such as ArcGIS before input into EMS-HAP. Often spatial allocation factors must be developed for the grids.

In an effort to allow more utility of EMS-HAP and when spatial programs such as ArcGIS are not available, polygons of all the census tracts in the U.S. were created from the year 2000 census (U.S.EPA 2004b). Datasets of the tract vertices or corners were developed as ancillary files for EMS-HAP. This update allows a user to have the ASPEN-ready surrogate files for use in ISCST3 or AERMOD. In AERMOD, these are designated as AREAPOLY sources (U.S. EPA, 2004a).

For the year 2000, 2010, and 2020 AERMOD simulations, the census tract polygons were used as the area sources in AERMOD. Note that area sources in this case refer to non-point sources in AERMOD.

For 1990, datasets of the tract polygons had to be developed in the same manner as the year 2000 tracts. Polygons were only developed for the three counties of interest in the study.

4.3 Emissions Processing

4.3.1 Point Inventory

EMS-HAP processing began with the module PtDataProc, which performed some location and stack parameter quality assurance. Sources with stack parameters exceeding a range set by the user were changed to default parameters based on the MACT, SIC, or SCC code of the source, or by defaults set by the user. PtDataProc also converted the coordinates of the sources from latitude and longitude to UTM coordinates if necessary.

After PtDataProc, the inventories were processed through PtModelProc, which speciated and grouped pollutants. For benzene, which is an inert pollutant, the manipulation of the inventory was minimal. Processing then continued through PtTemporal where the annual emissions were temporally allocated by season (winter, spring, summer, and fall), day of week (weekday, Saturday, or Sunday), and hour of day (hours 1 through 24).

After PtTemporal, processing was completed in PtFinal_ISC where source groups were assigned and AERMOD-formatted input files were created. For the point sources, two groups were created: major and area & other. Emissions were binned to these groups based on the value of the variable SRC_TYPE. For complete details of the different modules the reader is referred to the EMS-HAP User's Guide (U.S. EPA, 2004b). A flowchart of the processing is shown in Figure 12.

4.3.2 Nonpoint Inventory

Nonpoint emissions processing began with the EMS-HAP module COPAX, which splits the county-level nonpoint inventory into an airport-related emissions and remaining county-level nonpoint emissions. The airport-related emissions were assigned to the airports in the modeling domain based on the SCC and airport allocation factors (Tables 8 and 9). Once the emissions were split using COPAX, the airport emissions were processed in EMS-HAP using the same modules as for the point inventory (Section 4.3.1). The remaining county-level nonpoint emissions were processed through two modules, CountyProc and CountyFinal_ISCST3.

In CountyProc, the county-level emissions were spatially allocated to census tracts by using spatial surrogates. The surrogates allocate a portion of the county-level emissions to each tract based on a spatial allocation factor (SAF) assigned to that tract. For a particular spatial surrogate, the allocation factor is the ratio of the census tract level value for that surrogate to the county total value for the surrogate. For example, gasoline stations are used as a surrogate. Each census tract contains a number of gasoline stations. For all the census tracts in a particular county, say Harris County, the gasoline stations are summed to give a county total number of gasoline stations. Then for each tract, the number of gasoline stations for the tract is divided by the number of gasoline stations in the county. The general form of the SAF calculation is

$$SAF_{i,j} = \frac{Surrogate_{i,j}}{\sum_{1}^{j} Surrogate_{i}} \qquad \text{(Eq. 2)}$$

where $SAF_{i,j}$ is the spatial allocation factor for surrogate i for jth tract of the county. In the inventory, each emissions record was assigned a spatial surrogate based on the SCC of the record via a surrogate-SCC cross-reference file.

As well as spatially allocating the emissions to census tracts, CountyProc performed similar functions as in PtModelProc and PtTemporal, speciation and grouping of pollutants and temporal allocation of annual emissions by season, day of week and hour of day. As with the point inventory, the speciation and grouping functions of CountyProc were minimal since the pollutant was benzene. CountyProc also grouped the emissions into the source groups as assigned in Section 4.2.3.

Once processing was completed in CountyProc, the emissions were processed through CountyFinal_ISCST3, which assigned the tract polygon vertices to the census tracts. These data were then processed in AERMOD as AERMOD area sources. Source elevations by census tract were also assigned in CountyFinal_ISCST3. Currently EMS-HAP uses an ancillary file that contains the elevation of the census tract centroid for each census tract in the U.S. CountyFinal_ISCST3 assigns that elevation to the census tract source if source elevations are to be used. If the user chooses the flat terrain option in EMS-HAP elevations will not be output. The source elevation option was not used for the 1990 EMS-HAP processing, as census tract elevations were not available for the 1990 census tracts. This was not an issue since AERMOD was to be run with the FLAT terrain option. Source elevations were assigned to the other years' sources but those elevations were not used in AERMOD, again because the FLAT terrain option was chosen. The flow of emissions processing is shown in Figure 13.

4.3.3 Nonroad Processing

The nonroad inventories were processed in similar fashion as the nonpoint inventories, processing through COPAX to separate the airport emissions from the other nonroad sources. Airport emissions were processed through the point source programs, and the remaining nonroad sources were processed through CountyProc and CountyFinal. Refer to Figure 13 for the flow of processing.

4.3.4 Onroad Processing

Onroad emission inventories were processed through two of the EMS-HAP modules, PtTemporal and PtFinal_ISCST3. The inventories were only processed through those modules because when the emissions are assigned to the onroad links, it is not necessary to process the emissions through PtDataProc and PtModelProc. The emissions were processed through PtTemporal to create the 288 hourly emissions for the seasonal, day of week, and hourly emissions that are processed in AERMOD. Initially, the onroad emissions were by season and hourly, but not day of week. Finally, emissions were processed through PtFinal_ISCST3 to create the AERMOD input files. Figure 14 shows the steps in the onroad processing.

4.4 EMS-HAP Post-processing

After EMS-HAP processing, there was some post-processing of the emissions that was necessary for the building dimensions data for point sources and gas deposition data for all sources. For the building dimensions, EMS-HAP output building height and building width, as required by ISCST3. AERMOD also requires a building length and x and y adjustments. For this study the building length was set to 1.5 times the building width and the x and y adjustments were set to zero. Gas deposition parameters were needed for the deposition algorithms for AERMOD from Wesely et al (2002). These parameters are shown in Table 10 and are used for all inventories (i.e., point, nonpoint, nonroad, and onroad).

AERMOD also allows for the treatment of sources as urban or rural sources. After running EMS-HAP, a list of the urban sources, by source identification, was created for each inventory. For the nonpoint and nonroad inventories, the airport emissions were treated as rural sources and for onroad, all onroad links were treated as rural sources. Each point source was assigned the urban/rural designation of the census tract chosen to be the closest to the source during EMS-HAP processing. For the nonpoint and nonroad sources, since the tracts themselves were the sources, the source just took the urban/rural designation of the assigned census tract.

4.5 Emissions Summaries

Figure 15 shows the processed emissions by source group, by year, and by scenario (i.e., CAA and NON-CAA). Emissions are also listed in Table 11.

As can be seen from Table 11 and Figure 15, emissions for the NON-CAA inventories were projected to increase over the modeled timeframe relative to 1990, while the CAA inventories showed significantly lower emissions compared to the 1990 baseline, especially for mobile and stationary emissions. Table 12 summarizes the percent differences between the future year emissions and 1990 for the NON-CAA and CAA. For major sources, the CAA inventories were projected to have decreased 73%, 72%, and 67% for 2000, 2010, and 2020, respectively, from 1990 levels. Area & other sources decreased more than 80% for each of these scenarios. Onroad inventories were projected to have decreased by 68% to 88% over the same time scenarios. Nonroad differences also showed significant, although somewhat smaller, decreases over time. Across all sources, the total emissions initially decreased by 70% for 2000, but were estimated to remain fairly steady for 2010 (77%) and 2020 (76%).

For the NON-CAA inventories, the stationary emissions increased with time, but onroad mobile emissions were actually lower for the future years than for 1990; this result is likely due to the turnover of the vehicle fleet, increasing the number of vehicles with emission controls required as part of pre-1990 CAA requirements. The nonroad emissions grew with time as can be seen in the Tables 11 and 12 and Figure 15. Overall emissions grew slowly but by 2020 had increased by 29%.

Differences between the NON-CAA and CAA inventories for each year can also be seen in Table 12. For the major source emissions, the NON-CAA inventories were over 300% more than the CAA emissions. For area & other sources the differences were over 500%. Onroad mobile emissions differences increased with time from 102% to 606%. Nonroad emissions differences were also relatively small in 2000 (56%) and increased to 243% by 2020. Overall emissions differed by 251% in 2000 to 430% by 2020.

Figures 16 through 22 show breakdowns of emissions within each source category for each year and inventory scenario. For major and area & other, individual categories that comprised 90% of the emissions are shown, with other sources being around 10%. Note that the 1990 emissions inventory included "Chemical Manufacturing" in the nonpoint (or area source) inventory, while the 2000, 2010 and 2020 inventories classified this source category in the point (or major source) inventory. The figures depicting CAA emissions include Chemical Manufacturing in the point inventory, while the figures depicting NON-CAA emissions include this source category in the nonpoint inventory. The onroad inventories showed the contributions of gasoline and diesel emissions by road type (local or interstate roads). Nonroad emissions are shown by fuel type (2-stroke gasoline, 4-stroke gasoline, diesel, residual fuels, and aircraft). For 1990 and the NON-CAA inventories, the major and area & other inventories were dominated by the same categories, while the CAA inventories were dominated by a different set of categories. It also appeared that with time, the stationary emissions (major and area & other) became more dominant over the mobile categories.

Figure 23 shows the spatial distribution of the 1990 emissions for point, total allocated airport emissions (nonpoint plus nonroad) and tract allocated emissions (nonpoint plus nonroad). The larger point sources were located in eastern Harris County and the highest airport emissions were located at the two main airports in the region, George Bush and Hobby. The highest tract level emissions were located in Galveston County. These emissions were most likely from activities associated with the port in Galveston Bay near Houston. Figure 24 shows the 1990 onroad link emissions. The main interstates into the Houston area can be seen. However the emissions are so dense that it is difficult to discern emission patterns.

5 AERMOD

5.1 AERMOD Version

The version of AERMOD used in this study is Version 04300, the same version available on the EPA Support Center for Regulatory Atmospheric Modeling (SCRAM) website. There were some modifications made to the code:

- changes in model source code to allow execution on Linux system
- changes to allow area sources to run approximately 30% faster
- changes to the parameter, NVMAX, maximum number of AREAPOLY polygon vertices, in the module MAIN1, from 20 to 500 to allow for the use of the census tracts as AREAPOLY sources for the nonpoint and nonroad sources.

5.1.1 Control options

Within the AERMOD runstream file (the file used to control the execution of the model), several control options were implemented. These are shown in Table 13. For the AERMOD runs in this study, the TOXICS option was used. The TOXICS option is a non-regulatory option that specifies certain options to use in AERMOD such the SCIM option (Sampled Chronological Input Model), which samples meteorological data at specified intervals. However, in the AERMOD runs for this study, the SCIM option was not used. The TOXICS option also uses area optimization on area sources to decrease runtime.

The other control options listed in Table 13 include information about the urban population, seasons, land use, and other options. For details of these options see the AERMOD User's Guide (2004a).

5.1.2 AERMOD runs and post-processing

To decrease runtime and to take advantage of computer resources, the receptors were split into four sets for 1990 and eight sets for the other years. For each model run, hourly, daily, and annual average concentration output files were generated for each source category (major, area & other (excluding gasoline stations), gasoline stations, onroad, nonroad, and total). Once AERMOD runs were completed for each year, the annual average source category output files were read into SAS and concatenated into one dataset for later use in mapping and statistical calculations. Also gasoline stations and the other area & other concentrations were added together for a total area & other concentration at each receptor. Also, for 1990, several receptors were not included in analyses because they were located in census tracts in which the distance between the receptor and census tract was less than 100 meters. These were often receptors in very small census tracts, triangular in shape, or elongated in shape.

5.2 Results

Following are results for the seven modeling scenarios and model-to-monitor comparisons. The results are referred by the nomenclature given in Section 1.

5.2.1 Comparisons of scenarios

Results for the AERMOD simulations for each of the seven modeling scenarios are shown in Table 14 and in Figures 25 through 30. Table 14 lists the stationary, mobile, and total annual average concentrations by county for each scenario. Background is also listed only once since the same value is used for all scenarios. As expected from the emissions trends, concentrations increase between the CAA case and NON-CAA case for each year. Background contributions were not negligible and were sometimes larger than the modeled stationary or mobile concentrations.

Figure 25 shows the spatial distribution of total concentrations (all sources and background) for 1990 by receptor. The highest concentrations were located in Galveston County on Galveston Island. A high concentration was also located in east-central Harris County.

Figures 26 through 28 show the ratio of total NON-CAA concentrations to CAA concentration at each receptor for 2000, 2010, and 2020 respectively. Concentrations differed mostly within central Houston, with the NON-CAA concentrations 20 to almost 60 times larger than the CAA concentrations for each year. There are also pockets of high ratios in Galveston County and southern Brazoria County.

Figure 29 shows a comparison of the concentration distributions for all the modeling scenarios. For major sources, the concentrations steadily rose with the NON-CAA scenarios for each modeling period after 1990. These modeled concentrations were much higher than the CAA concentrations with median estimates two to three times larger for the NON-CAA distributions than the CAA inventories. For the area & other concentrations, the distributions for the NON-CAA distributions remained fairly steady from year to year with the same trend seen for the CAA distributions. For the onroad distributions, the 2000 distributions exhibited more overlap than the other years, with a decrease in the CAA concentrations with time (2000 to 2020). The overlap of 2000 was also true with the nonroad distributions with a decrease in CAA concentrations to 2010. From 2010 to 2020, concentrations remained fairly steady.

Figure 30 shows the concentration distributions for total concentrations without background and with background included. Separate distribution plots were made to show how background could potentially mask differences since background concentrations were fairly large. The distributions show that background did not affect the overall trending of the concentrations.

Table 15 lists the percent differences between 1990 and each CAA and NON-CAA inventory as well as the differences between the NON-CAA and CAA inventories for each year. For major sources, onroad mobile sources, and nonroad mobile sources, the concentration trends agreed well with the emission differences. However, the differences for area & other concentrations differed from the emissions differences for area & other in that while emissions increased from 1990 for the NON-CAA cases, the concentrations actually decreased. Possible reasons for the differences could be:

1. Differences between 1990 and 2000 meteorology. Annual averages for variables such as wind speed, wind direction, and stability may be similar between the two years; however, on an hourly or daily basis there may be differences large enough to change how the emissions are dispersed.

2. Differences in receptor locations. The receptors were based on census block groups for 1990 or 2000. While the emission sources were in the same locations for 1990 and the NON-CAA inventories, the receptors were distributed differently relative to the sources for 1990 and the other years. A subtle change in source-receptor distance could lead to differing concentrations.

3. Differences in land use for urban-rural designation or spatial allocation. Different definitions were used for urban designation (See Section 2.5) for 1990 and the other years. Also, different spatial surrogates and different census tract sources were used between 1990 and the other years, which could result in a different spatial distribution of emissions.

The factors described above may have been large enough to counter the increase in emissions. For the total concentrations and total emissions, trends also are opposite for the NON-CAA cases. This is due to the large contribution of the area & other sources to the total concentrations as described below.

Figure 31 shows the percent contribution by source category to the annual average total concentrations across the domain. For 1990 and the NON-CAA concentrations, the background contribution was near 25% while for the CAA concentrations, the background contribution was near 60 to 66%. The increase in contribution was due to the decrease in concentrations of the CAA cases because of lower emissions, while the same background was applied to all scenarios. For 1990 and the NON-CAA cases, the dominant source category was the area & other category, initially at 54% for 1990 and 41% or 42% for the other years. For the CAA cases, other than background contributions, the major and area & other categories were usually very close together percentage wise.

5.2.2 Model-to-monitor comparisons

For model comparisons, the locations of 15 monitors were used as receptor inputs in the AERMOD simulations for the year 2000 (with CAA control) and 24-hour average concentrations were calculated at these monitor locations. The monitor locations and the annual average observed concentrations are shown in Figure 10. The highest observed concentrations were located in southwestern Harris County and southwestern Galveston County. Figure 32 shows the distributions of all the daily monitored and modeled concentrations. For daily modeled concentrations, only days with monitor observations were included in calculations. It can be seen that on a point-to-point basis the mean modeled concentrations are generally lower than the observed concentrations when background values were not included in the total modeled concentrations. Including background concentrations led to better agreement with the observed concentrations. Also, the distribution of modeled concentrations shows less variation between the 5th and 95th percentiles than when background is excluded. This is due to the large contribution of background to the total modeled concentrations and its nonuniformity within the counties. (Refer to Section 2.7 for how background concentrations were estimated).

Figure 33 shows the model-to-monitor ratios based on the daily average concentrations for the monitors and model (with and without background). Ratios between 0.5 and 2 generally refer to good agreement between the model and the observed concentrations. Without background, the model to monitor ratios are generally lower than 0.5 but including background improves the model comparison as most ratios are between the 0.5 and 2 range.

6 HAPEM6

This section describes the Hazardous Air Pollutant Exposure Model, version six (HAPEM6), the creation of the air quality files, the processing of the model, and the results.

6.1 HAPEM6 Description

Exposure modeling was done using the HAPEM6 model (U.S. EPA, 2007), which is based on the HAPEM5 model (U.S. EPA, 2005). One of the main differences between HAPEM6 and HAPEM5 is that HAPEM6 accounts for near-roadway concentrations.

The HAPEM6 exposure model used in this assessment is the most recent version in a series of models that the EPA has used to model population exposures and risks at the urban and national scale in a number of assessments (U.S. EPA, 1993; U.S EPA, 1999; U.S. EPA, 2002c). HAPEM6 is designed to assess average long-term inhalation exposures of the general population, or a specific sub-population, over spatial scales ranging from urban to national. HAPEM6 uses the general approach of tracking representatives of specified demographic groups as they move among indoor and outdoor microenvironments and among geographic locations. The estimated pollutant concentrations in each microenvironment visited are combined into a time-weighted average concentration, which is assigned to members of the demographic group. HAPEM6 calculates 30 replicates with different exposures for each demographic group. These data can be used to develop a distribution of exposures for the entire U. S. population.

HAPEM6 uses five primary sources of information: population data, population activity data, air quality data, roadway locations, and microenvironmental data. The population data used is obtained from the U.S. Census. Two kinds of activity data are used: activity pattern data and commuting pattern data. The activity pattern data quantify the amount of time individuals spend in a variety of microenvironments and come from EPA's Consolidated Human Activity Database (CHAD) (Glen et al., 1997). The commuting data contained in the HAPEM6 default file were derived from the year 2000 U.S. Census, and include the number of residents of each tract that work in that tract and every other U.S. Census tract, as well as data on commuting times and distances. The air quality data come from AERMOD (after background has been added). The road locations are determined from geographic information system files from the U.S. Census. The microenvironmental data consist of factors that estimate air toxic concentrations in specific microenvironments, based on penetration of outdoor air into the microenvironment, proximity of the microenvironment to the emission source, and emission sources within the microenvironment. These factors vary among pollutants (Long et al., 2004).

New to HAPEM6 are algorithms that account for the gradient in concentrations of primary (directly emitted) mobile source air toxics within 200 meters of major roadways (U.S. EPA, 2007). HAPEM6 adjusts ambient concentrations generated by ASPEN for each census tract using concentration gradients developed using the CALPUFF dispersion model (Cohen et al., 2005). For locations within 75 meters and from 75 to 200 meters from major roads, ambient concentrations are adjusted upward, while locations further from major roadways are adjusted downward. These adjustments are consistent with results from prior modeling studies that explicitly accounted for concentration gradients around major roads within census tracts (Pratt et al., 2004). These adjusted concentrations are then employed in microenvironmental concentration calculations.

HAPEM6 has a number of other technical improvements over the previous version of HAPEM5. These improvements, along with other details of the model, are described in the HAPEM6 User's Guide (U.S. EPA 2007). The HAPEM6 runs used year 2000 census data. Average lifetime exposure for an individual in a census tract was calculated from data for individual demographic groups using a post-processing routine. We estimated the contributions to ambient concentrations for the following source sectors: major, area and other, onroad, nonroad, and background.

6.2 Creation of air quality files

Ambient air quality concentrations were input into HAPEM6 in eight three-hour blocks for each source category and one annual value for background at each census tract or block group. To calculate the eight three-hour concentrations at each block group for each source group, the hourly concentrations for each receptor, or block group, were assigned a time block designation. Block 1 for hours 1 through 3, block 2 for hours 4 through 6, block 3 for hours 7 through 9, block 4 for hours 10 through 12 and so forth. Also, the gasoline station concentrations were added to the other area & other concentrations to yield a total area & other concentration for each hour and receptor. As with the annual average concentration analyses, the receptors located less than 100 meters from a census tract were dropped from consideration. Concentrations for each source group were averaged across all hours of the year for each time block (i.e., all hours 1 through 3 for time block 1) for each receptor.

Concentrations for hours with missing meteorological or calm winds were not included in the calculations. This is because AERMOD assigned a value of zero to concentrations for hours with missing meteorological data or calm winds. Within AERMOD, those hours were not included in the annual average concentrations. So that the average of the eight-hour averages $[(X_1+X_2+X_3...+X_8)/8]$ equaled the annual average concentration for each receptor and source group, those hours were not included in the calculation of the HAPEM6 input concentrations. For any AERMOD simulation, a list of the hours with missing or calm winds can be found in the ERRORS.OUT file. After calculating the eight three-hour averages, the county-specific background was then added to the data file.

For 1990, the census tracts from the 1990 census could not be used for HAPEM6, because the commuting file in the HAPEM6 model was based on the year 2000 census tracts. Therefore, in order to process the 1990 AERMOD results in HAPEM6, the 1990 block groups were assigned the year 2000 census tracts in which they were located. This was done by plotting the receptor coordinates and year 2000 census tracts in ArcGIS and assigning the 1990 block groups to the year 2000 tracts where they overlapped. For all other years, this step was not necessary as they used the year 2000 block groups as receptors.

Once the eight three-hour concentrations were calculated for each block group and background assigned, the concentrations were written to a text file ready for input into HAPEM6. The format of the file is State/County FIPS, census tract, background, and then eight three hourly average concentrations for major, area & other, onroad, and nonroad, in that order. A second file was created just for gasoline stations, with gasoline stations taking the place of area & other and all other concentrations set to zero, including background.

6.3 HAPEM6 Processing

The HAPEM6 model consists of 5 modules, to be run in order: 1) DURAV6, 2) INDEXPOP6, 3) COMMUTE6, 4) AIRQUAL6 and 5) HAPEM6. The first three modules (DURAV6, INDEXPOP6, and COMMUTE6) are used to determine human activity patterns and demographic information appropriate to the study area. The AIRQUAL module processes the ambient concentration data from AERMOD into the proper model format. Finally, the last module, HAPEM6 calculates inhalation exposure concentrations at each census tract. Further details on HAPEM6 can be found in the User's Guide (U.S. EPA, 2007).

The outputs of HAPEM6 are exposure concentration files among several demographic groups for each census tract. A series of post-processing FORTRAN programs calculates the distribution of exposure concentrations for each source category at each census tract. The final output concentrations are the 1st, 5th, 10th, 25th, 75th, 90th, 95th, and 99th percentiles, average, and median concentration for each source category and census tract.

6.4 HAPEM6 Results

County average HAPEM6 exposure concentrations are shown in Table 16. The averages were calculated from the median exposure concentrations at each tract. Note that for HAPEM6, the background differs for each modeling scenario, even though the same background was input into HAPEM6 for each scenario. Since the same background was input for each scenario, it would stand to reason that the exposure background concentrations would be the same for each scenario. The difference is due to post-processing of the raw HAPEM6 results in order to calculate the exposure concentrations. The source category and background concentrations are adjusted by the median total concentration for each tract. Therefore, a change in one source category into HAPEM6, with all others remaining the same, can lead to differences in exposure concentrations for all categories, because the total concentration changes. As with the AERMOD concentrations, the HAPEM6 exposure concentrations increased with the NON-CAA scenarios when compared to the CAA scenarios for each year.

Figure 34 shows the spatial distribution of the 1990 HAPEM6 exposure total concentrations (all sources and background) at the tract level. Note that the year 2000 tracts were used for 1990. Tracts in white are those tracts in which no 1990 block group centroid was located. As with the AERMOD concentrations, the highest concentrations were found on Galveston Island. Figures 34 through 36 are the ratio of the NON-CAA to CAA HAPEM6 total concentrations at each tract for each year, 2000, 2010, and 2020 respectively. A similar distribution was seen as with the

AERMOD concentrations: higher ratios in Houston, the coast of Galveston County on Galveston Bay, and southern Brazoria, with similar magnitudes.

Figure 38 shows the distributions of the CAA and NON-CAA HAPEM6 concentrations for major, area & other, onroad, and nonroad source categories. The HAPEM6 concentrations exhibited similar behavior as the AERMOD distributions. For HAPEM6, there appeared to be less overlap for the area & other concentrations as for the AERMOD distributions. This was probably due to the commuting within HAPEM6 as people commuted to or near a large source in the area & other category.

Figure 39 shows the distributions of background, total concentrations without background, and total concentrations with background. Unlike the other source categories, the background concentrations actually decreased with the NON-CAA scenarios. This can be explained by the background remaining fairly constant between the CAA and NON-CAA scenarios, due to same background being used for all scenarios, but total concentrations increased from the CAA to NON-CAA scenarios due to increased emissions. When dividing the fairly constant (between scenarios) background by the total median concentration, the adjusted background for the NON-CAA scenarios were less than the CAA scenarios. The total (with and without background) exhibited similar behavior as with the AERMOD total concentrations.

Figure 40 shows the percent contribution by source category to the average HAPEM6 concentrations. As with the AERMOD concentrations, the background contribution was largest for the CAA concentrations and stationary contributions were larger than for mobile contributions except for 2000 CAA. For 1990 and the NON-CAA concentrations, the area & other concentrations were the larger contributors for each year.

Table 17 shows the percent differences between 1990 and each future year domain average HAPEM6 concentration for each source category as well as differences between NON-CAA and CAA concentrations for each year. Generally, the HAPEM6 percent differences followed those of the AERMOD differences with the exception of the difference between 2000 CAA and 1990 nonroad concentrations, which differed by about 2%. This agreed with the distributions of nonroad concentrations shown in Figure 38.

Figure 41 shows the distribution of HAPEM6 to AERMOD concentrations for each source category and Table 18 shows the average ratio for each source category. The ratios can be used to show how important commuting is in HAPEM6. For these ratios, the block group concentrations were averaged to the census tract level. For 1990, this means the year 2000 tracts. For major sources, the ratio was mostly less than 1.0, meaning commuting in HAPEM6 was actually moving people away from emission sources in their home tracts. Area & other ratios were move above 1.0, meaning people are being moved toward higher emissions. For onroad, almost all ratios were above 1.0 with nonroad between 0.8 and 2.

Figure 42 shows the ratio distributions for background and total concentrations (without background and with background). Background ratios were below 1 and the differences between the CAA and NON-CAA scenarios can be explained as above for Figure 41. When background was not included in total concentrations, the distributions of the ratios between the CAA and NON-CAA scenarios were very similar.   When background was included, the ratios were higher for the NON-CAA scenarios.

7 Limitations

Limitations of this study include the following:

1. Data availability/reliability. Emissions data can have uncertainties in emissions magnitude and locations. The user and EMS-HAP quality assurance procedures address these issues. Meteorological data, terrain inputs, and site selection can also contribute to model results uncertainties. Quality assurance should be done by the user and preprocessing programs, such as AERMET for the meteorological data.

2. Emissions domain. The emissions covered a limited domain of three counties. Any model receptors located near county boundaries may have underestimated concentrations when compared against monitor concentrations. Model receptors near the county boundaries are not affected by neighboring county emissions whereas monitors are affected.

3. Computational resources. Computational resources can limit the number of receptors and averaging period of the model. Increasing receptors and averaging period can increase model runtimes and increase use of computational resources.

4. Model limitations. AERMOD model formulations contribute to uncertainty. See the AERMOD User's Guide for limitations (U.S. EPA, 2004a).

5. During post-processing of HAPEM6 exposure concentrations, the post-processing code adjusts the source category concentrations by the total median concentration at the tract. Due to this calculation, changes can occur to exposure concentrations to a source category between modeling scenarios, even if the input concentrations into HAPEM6 did not differ between scenarios, e.g. background concentrations.

6. As stated in 2, only three counties were involved in the study. This may affect the commuting flow in HAPEM6 as some tracts may have commutes outside the three counties into neighboring counties.

7. Differences between 1990 and future year benzene concentrations cannot be solely explained by emission differences between these model years. This is due to different meteorology, spatial surrogates, and different receptor spatial distributions used for 1990 and the other years. To see the "true" effects of the Clean Air Act controls, all years can be run with the same meteorology, surrogates, and receptors to eliminate differences from those factors.

8. The modeling does not account for impacts of demographic shifts that are likely to occur in the future. These changes in demographics will affect our estimates of exposure.

9. The modeling does not account for indoor sources or non-inhalation pathways of exposure.

10. A key limitation is using 1999 “background” levels to account for mid-range to long-range transport. However, since background is related to emissions far away from receptors, these levels should decrease as those emissions decrease.

11. Use of surrogates to allocate nonroad and area source emissions to census tracts may distribute emissions to large areas, thus lowering emission density.

12. MOBILE6.2 underestimates cold start emissions for Tier 1 and later vehicles at cold temperatures.

13. Portable fuel containers, which are a significant source of benzene emissions, are not included in the inventories.

14. Modeling does not include recent revisions to EPA's NONROAD model which include new evaporative categories for tank permeation, hose permeation, hot soak, and running loss emissions, a revised methodology for calculating diurnal emissions, and improvements to allocation of emissions from recreational marine and construction equipment.

8 Summary and Conclusions

Emissions for 1990, 2000, 2010, and 2020 were processed through EMS-HAP, AERMOD, and HAPEM6 to yield ambient and exposure concentrations. For 2000, 2010, and 2020, emissions reflected controls by the 1990 Clean Air Act and without Clean Air Act controls. The 1990 inventory was considered a base inventory, as this was the year of the Clean Air Act Amendments. For all emission scenarios, nonpoint and nonroad county-level emissions were allocated to census tracts using spatial surrogates or either allocated to airports. Onroad emissions were allocated to links to refine the location of onroad emissions, and point sources were modeled according to their locations. Census tracts were modeled in AERMOD as polygons, airports as area sources, onroad emissions as elongated area sources, and point sources as points. From the results several conclusions can be drawn:

- Without Clean Air Act controls, total emissions would have increased from 1990 to 2020 increased by 29%.

- With Clean Air Act controls, total emissions are projected to decrease from 1990 to 2020 by 76%.

- Among 2000, 2010, and 2020, NON-CAA emissions would have increased an average of 353% with a maximum of 430% in 2020.

- Without Clean Air Act controls, total (all sources) annual average AERMOD and HAPEM6 concentrations would have decreased by 8% from 1990 to 2020. The decreases in total concentrations were due to decreases in the area & other concentrations, as noted in Section 5.2.

- Without With Clean Air Act controls, total (all sources) annual average AERMOD and HAPEM6 concentrations would have decreased by 68% from 1990 to 2020.

- Without Clean Air Act controls in 2000, 2010, and 2020, AERMOD concentrations from all sources would have increased an average of 159% with a maximum of 187% in 2020. Similarly, for the same scenarios, annual average HAPEM6 concentrations increased an average of 154% with a maximum of 185% for 2020.

- Among all cases, 1990, and all CAA and all NON-CAA cases, total (all sources) annual average HAPEM6 concentrations were around 90% of the total AERMOD concentrations. This result implied that HAPEM6 modeled people commuting from census tract with higher concentrations to tracts with lower concentrations.

- Model validation with observed monitor concentrations showed the importance of inclusion of background concentrations with model concentrations as model agreement with the monitors increased when background was included with the modeled concentrations.

- Given the differences in emission trends and modeled concentration trends, it may be useful to run 1990 in EMS-HAP and AERMOD with year 2000 spatial allocation factors, 2000 meteorology, and 2000 census block group receptors to see if those same trends seen with concentrations emerge when the only differing variables are the emissions magnitudes.

9 References

Battelle, 2003: Estimate background concentrations for the National-Scale Air Toxics Assessment. Technical Report. Contract No. 68-D-02-061. Work Assignment 1-03.

Cohen, J., R. Cook, C.R. Bailey, E. Carr, 2005. Relationship between motor vehicle emissions of hazardous pollutants, roadway proximity, and ambient concentrations in Portland, Oregon. Environ Modeling & Software 20: 7-12.

Cook, R., E. Glover, 2002. Technical Description of the Toxics Module for MOBILE6.2 and Guidance on Its Use for Emission Inventory Preparation. U.S. EPA, Office of Transportation and Air Quality, Ann Arbor, MI. Report No., EPA420-R-02-011. http://www.epa.gov/otaq/m6.htm

E.H. Pechan and Associates, 2006. Section 812 Clean Air Act Cost-Benefit Study: Air Toxics Case Study Benzene Emissions Reductions in Houston. Prepared for U. S. EPA, Emission Factor and Inventory Group, Emissions Monitoring and Analysis Division, Office of Air Quality Planning and Standards, EPA, February 2006.

E.H. Pechan, Associates. Personal Communication. 22 March 2006

Glen, G., Y. Lakkadi, J. A. Tippett, M. del Valle-Torres. 1997. Development of NERL/CHAD: The National Exposure Research Laboratory Consolidated Human Activity Database. Prepared by ManTech Environmental Technology, Inc. EPA Contract No. 68-D5-0049.

Long, T, T. Johnson, J. Laurensen, A. Rosenbaum 2004. Development of Penetration and Proximity Microenvironment Factor Distributions for the HAPEM5 in Support of the 1999 National-Scale Air Toxics Assessment (NATA). Memorandum from TRJ Consulting and ICF Consulting, Inc. to Ted Palma, U. S. EPA, Office of Air Quality Planning and Standards, RTP, NC, April 5, 2004. http://www.epa.gov/ttn/fera/human_hapem.html

Pratt, G. C., C. Y. Wu, D. Bock, et al. (2004) Comparing air dispersion model predictions with measured concentrations of VOCs in urban communities. Environ. Sci. Technol. 38: 1949-1959.

U. S. Environmental Protection Agency, 1993. Motor Vehicle-Related Air Toxics Study. Office of Mobile Sources, Ann Arbor, MI. Report No. EPA 420-R-93-005. http://www.epa.gov/otaq/regs/toxics/tox_archive.htm

U.S. Environmental Protection Agency, 1995. User's Guide for the Industrial Source Complex (ISC3) Dispersion Models: Volume I – User Instructions. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454-/B-95-003a. http://www.epa.gov/scram001/userg/regmod/isc3v1.pdf

U.S Environmental Protection Agency, 1999. Analysis of the Impacts of Control Programs on Motor Vehicle Toxics Emissions and Exposure in Urban Areas and Nationwide. Prepared for U. S. EPA, Office of Transportation and Air Quality, by Sierra Research, Inc., and Radian International Corporation/Eastern Research Group. Report No. EPA 420 –R-99-029/030. http://www.epa.gov/otaq/regs/toxics/tox_archive.htm

U. S. Environmental Protection Agency, 2000. User's Guide: Assessment System for Population Exposure Nationwide (ASPEN, Version 1.1) Model. Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454-R-00-017. http://www.epa.gov/scram001/userg/other/aspenug.pdf

U. S. Environmental Protection Agency, 2002a. Example Application of Modeling Toxic Air Pollutants in Urban Areas, Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/R-02-003. http://www.epa.gov/scram001/guidance/guide/uatexample.pdf

U. S. Environmental Protection Agency, 2002b. User's Guide for the Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP, Version 2.0), Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-02-001. http://www.epa.gov/scram001/userg/other/emshapv2ug.pdf

U. S. EPA. 2002c. 1996 National-Scale Air Toxics Assessment. http://www.epa.gov/ttn/atw/nata/

U.S. Environmental Protection Agency, 2004a. User's Guide for the AMS/EPA Regulatory Model – AERMOD, Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-001. http://www.epa.gov/scram001/7thconf/aermod/aermodugb.pdf

U. S. Environmental Protection Agency, 2004b. User's Guide for the Emissions Modeling System for Hazardous Air Pollutants (EMS-HAP, Version 3.0), Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-006. http://www.epa.gov/scram001/dispersion_related.htm#ems-hap

U.S. Environmental Protection Agency, 2004c. User's Guide for the AERMOD Meteorological Preprocessor (AERMET), Office of Air Quality Planning and Standards, Research Triangle Park, NC, Report No. EPA-454/B-03-002. U.S. http://www.epa.gov/scram001/7thconf/aermod/aermetugb.pdf

U.S. Environmental Protection Agency, 2004d. NONROAD2004 Model, Office of Transportation and Air Quality, Ann Arbor, MI. http://www.epa.gov/otaq/nonrdmdl.htm.

U. S. Environmental Protection Agency, 2005: The HAPEM5 User's Guide - Hazardous Air Pollutant Exposure Model, Version 5; March 2005, Available at: http://www.epa.gov/ttn/fera/hapem5/hapem5_guide.pdf ; RTP, NC

U. S. Environmental Protection Agency, 2007: The HAPEM6 User's Guide - Hazardous Air Pollutant Exposure Model, Version 6; January 2007, Available at: http://www.epa.gov/ttn/fera/human_hapem.html; RTP, NC

Wesely, M.L., P.V. Doskey., and J.D. Shannon, 2002. Deposition Parameterizations for the Industrial Source Complex (ISC3) Model, Draft, ANL/ER/TM-nn, DOE/xx-nnnn, Environmental Research Division, Argonne National Laboratory, Argonne, IL. Available on http://www.epa.gov/scram001

Table 1. Year of data used in EMS-HAP and AERMOD for each inventory scenario.

Data	Inventory						
	1990	2000 CAA	2000 NON-CAA	2010 CAA	2010 NON-CAA	2020 CAA	2020 NON-CAA
Receptors	1990	2000	2000	2000	2000	2000	2000
Urban/rural classification	1990	2000	2000	2000	2000	2000	2000
Meteorological inputs	1990	2000	2000	2000	2000	2000	2000
Spatial surrogates	1990	2000	2000	2000	2000	2000	2000

Table 2. Surface and upper air meteorological variables output from AERMET for use in AERMOD.

Surface	Upper Air
Sensible Heath flux (W m^{-2})	Measurement height (up to 9)
Surface Friction Velocity u* (m s^{-1})	Indicator flag (indicates last level for hour)
Convective Velocity Scale w* (m s^{-1})	Wind speed (m s^{-1})
Potential temperature lapse rate above mixing height (Km^{-1})	Wind direction (degrees)
Convective mixing height (m)	Temperature (˚C)
Mechanical mixing height (m)	σ_a: standard deviation of lateral wind direction (deg)
Monin-Obukhov length (m)	σ_a: standard deviation of vertical wind speed (m s^{-1})
Z_o: surface roughness length (m)	
Daytime Bowen ratio	
Noon time albedo	
Surface wind speed (m s^{-1})	
Wind direction (deg)	
Height of surface wind (10m preferred)	
Precipitation code (integer)	
Precipitation amount (mm hr^{-1})	
Relative humidity (%)	
Surface pressure (mb)	
Cloud cover (tenths)	
Ambient surface temperature (K)	

Table 3. Comparisons of wind speed, wind direction, temperature, and annual rainfall for IAH 1990, HOU 2000, and climatology. 1990 and 2000 values based on post AERMET data.

Variable	Climatology	IAH 1990	HOU 2000
Wind speed (m s^{-1})[1]	3.6	3.7	3.7
Wind direction (degrees) [1]	153	151	142
Average minimum daily temperature (˚C) [1]	15.3	15.8	17.0
Average maximum daily temperature (˚C)[1]	25.9	26.7	26.5
Annual rainfall (mm) [2]	1,170	1,025	1,102
1. Climatology based on 1984-1992. 2. Climatology based on 1961-1990			

Table 4. County specific background values used in study.

FIPS	County	Concentration (µg per m^3)
48039	Brazoria	0.36
48167	Galveston	0.40
48201	Harris	0.46

Table 5. SCC codes of sources characterized as volume sources.

SCC	Description
28888801	Internal Combustion Engines; Fugitive Emissions; Other Not Classified; Specify in Comments
30113227	Industrial Processes; Chemical Manufacturing; Organic Acid Manufacturing; Fugitive Emissions
30115680	Industrial Processes; Chemical Manufacturing; Cumene; Fugitive Emissions
30116980	Industrial Processes; Chemical Manufacturing; Ethyl Benzene; Fugitive Emissions
30119749	Industrial Processes; Chemical Manufacturing; Butylene, Ethylene, Propylene, Olefin Production; Ethylene: Fugitive Emissions
30120580	Industrial Processes; Chemical Manufacturing; Propylene Oxide; Fugitive Emissions
30120680	Industrial Processes; Chemical Manufacturing; Styrene; Fugitive Emissions
30125880	Industrial Processes; Chemical Manufacturing; Benzene/Toluene/Aromatics/Xylenes; Aromatics: Fugitive Emissions
30130380	Industrial Processes; Chemical Manufacturing; Allyl Chloride; Fugitive Emissions
30130580	Industrial Processes; Chemical Manufacturing; Epichlorohydrin; Fugitive Emissions
30180001	Industrial Processes; Chemical Manufacturing; General Processes; Fugitive Leaks
30188801	Industrial Processes; Chemical Manufacturing; Fugitive Emissions; Specify in Comments Field
30188802	Industrial Processes; Chemical Manufacturing; Fugitive Emissions; Specify in Comments Field
30188803	Industrial Processes; Chemical Manufacturing; Fugitive Emissions; Specify in Comments Field
30188805	Industrial Processes; Chemical Manufacturing; Fugitive Emissions; Specify in Comments Field
30600801	Industrial Processes; Petroleum Industry; Fugitive Emissions; Pipeline Valves and Flanges
30600805	Industrial Processes; Petroleum Industry; Fugitive Emissions; Miscellaneous: Sampling/Non-Asphalt Blowing/Purging/etc.
30600816	Industrial Processes; Petroleum Industry; Fugitive Emissions; Flanges: All Streams
30688801	Industrial Processes; Petroleum Industry; Fugitive Emissions; Specify in Comments Field
30988801	Industrial Processes; Fabricated Metal Products; Fugitive Emissions; Specify in Comments Field
31000207	Industrial Processes; Oil and Gas Production; Natural Gas Production; Valves: Fugitive Emissions
31000220	Industrial Processes; Oil and Gas Production; Natural Gas Production; All Equipt Leak Fugitives (Valves, Flanges, Connections, Seals, Drains
31088801	Industrial Processes; Oil and Gas Production; Fugitive Emissions; Specify in Comments Field
31088803	Industrial Processes; Oil and Gas Production; Fugitive Emissions; Specify in Comments Field
31088811	Industrial Processes; Oil and Gas Production; Fugitive Emissions; Fugitive Emissions
40288801	Petroleum and Solvent Evaporation; Surface Coating Operations; Fugitive Emissions; Specify in Comments Field
40388801	Petroleum and Solvent Evaporation; Petroleum Product Storage at Refineries; Fugitive Emissions; Specify in Comments Field
40688801	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Fugitive Emissions; Specify in Comments Field
40688802	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Fugitive Emissions; Specify in Comments Field
49000206	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Waste Solvent Recovery Operations; Fugitive Leaks
50300801	Waste Disposal; Solid Waste Disposal - Industrial; Treatment, Storage, Disposal/TSDF; Surface Impoundment: Fugitive Emissions

Table 6. Stacks with adjusted coordinates.

Site ID	Emissions release point ID	Sitename	Original longitude	Original latitude	New longitude	New latitude	Basis for change[#]
48039-NEI11295	1-49	DOW CHEMICAL CO. TEXAS OPERATIONS	-95.02072337	28.89581637	-95.37889	28.98611	Located outside listed county of site. Use average coordinates of stacks from same site
48039-NEI6519	1-63	SWEENY REFINERY PETROCHEM	-94.22985009	28.606079768	-95.71493742	29.044421399	Located outside listed county of site. Use average coordinates of stacks from same site
48167-NEI2TXT17872	01-138	Amoco Oil Co., Docks Nos. 37 and 38.	-97.4431	27.8108	-94.8897205	29.369445	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NEI6625	1-477	HOUSTON OLEFINS PLANT	-93.21688059	29.092470972	-95.25067286	29.702064476	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NEI6625	1-478	HOUSTON OLEFINS PLANT	-93.21688059	29.092470972	-95.25067286	29.702064476	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NEI7229	1-575	CHEVRON CHEMICAL CO	-91.21671407	28.660275149	-94.91737572	29.823302167	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NEI7229	1-576	CHEVRON CHEMICAL CO	-91.21671407	28.660275149	-94.91737572	29.823302167	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NE7741	1-594	BAYTOWN OLEFINS PLANT	-92.78986181	29.056455647	-95.005051	29.762074	Located outside listed county of site. Use average coordinates of stacks from same site
48201-NEI11119	1-363	DEER PARK SITE	-94.66997077	29.577514503	-95.12917	29.72139	Located outside listed county of site. Use average coordinates of stacks from same site

Table 7. Airport related SCC codes and assigned airport allocation factor type.

SCC	Description	Allocation factor type
2501080000[#]	Aviation Gasoline Distribution: Stage 1 & II	General Aviation
2501080050[#]	Aviation Gasoline Storage -Stage I	General Aviation
2501080100[#]	Aviation Gasoline Storage -Stage II	General Aviation
2265008000	Airport Support Equipment, Total, Off-highway 4-stroke	Commercial Aviation
2265008005	Airport Support Equipment, Off-highway 4-stroke	Commercial Aviation
2267008000	Airport Ground Support Equipment, All, LPG	Commercial Aviation
2267008005	Airport Ground Support Equipment, LPG	Commercial Aviation
2268008000	Airport Ground Support Equipment, CNG, All	Commercial Aviation
2270008000	Airport Service Equipment, Total, Off-highway Diesel	Commercial Aviation
2270008005	Airport Service Equipment, Airport Support Equipment, Off-highway Diesel	Commercial Aviation
2275000000	All Aircraft Types and Operations	Commercial Aviation
2275001000	Military Aircraft, Total	Military Aircraft
2275020000	Commercial Aircraft, Total	Commercial Aviation
2275050000	General Aircraft, Total	General Aviation
2275060000	Air Taxi, Total	Air Taxi
2275070000	Aircraft Auxiliary Power Units, Total	Commercial Aviation
2275900000	Aircraft Refueling: All Fuels, All Processes	Commercial Aviation

[#] nonpoint inventory

Table 8. Allocation factors by type by airport for 1990, 2000, 2010, and 2020.

Year	Allocation type	Airport								
		LBX	GLS	SPX	DWH	EFD	HOU	IAH	IWS	T41
1990	General	1.000	0.426	0.574	0.214	0.082	0.406	0.132	0.100	0.066
	Commercial	1.000	0.426	0.574	0	0.009	0.351	0.640	0	0
	Air Taxi	1.000	1.000	0	0.016	0.033	0.237	0.700	0.014	0
	Military	1.000	1.000	0	0.017	0.820	0.075	0.088	0	0
2000	General	1.000	0.925	0.075	0.272	0.133	0.314	0.081	0.096	0.105
	Commercial	1.000	0.925	0.075	2.23×10^{-6}	0.011	0.271	0.718	0	0
	Air Taxi	1.000	1.000	0	0.009	0.037	0.149	0.800	0.006	0
	Military	1.000	1.000	0	0.220	0.673	0.073	0.035	0	0
2010	General	1.000	0.956	0.044	0.282	0.164	0.273	0.055	0.102	0.123
	Commercial	1.000	0.956	0.044	0	0	0.316	0.684	0	0
	Air Taxi	1.000	1.000	0	0.007	0.008	0.137	0.845	0.002	0
	Military	1.000	1.000	0	0.058	0.925	0.010	0.007	0	0
2020	General	1.000	0.962	0.038	0.292	0.134	0.291	0.059	0.097	0.127
	Commercial	1.000	0.962	0.038	0	0	0.306	0.694	0	0
	Air Taxi	1.000	1.000	0	0.006	0.006	0.124	0.862	0.002	0
	Military	1.000	1.000	0	0.058	0.925	0.010	0.007	0	0

Table 9. Gasoline station SCC codes.

SCC	Description
2501060052	Storage and Transport; Petroleum and Petroleum Product Storage; Gasoline Service Stations; Stage 1: Splash Filling
2501060053	Storage and Transport; Petroleum and Petroleum Product Storage; Gasoline Service Stations; Stage 1: Balanced Submerged Filling
2501060100	Storage and Transport; Petroleum and Petroleum Product Storage; Gasoline Service Stations; Stage 2: Total
2501060201	Storage and Transport; Petroleum and Petroleum Product Storage; Gasoline Service Stations; Underground Tank: Breathing and Emptying

Table 10. Gas deposition parameters for benzene for AERMOD.

Parameter	Value
Diffusivity in air, D_a ($cm^2\ s^{-1}$)	0.08962
Diffusivity in water, D_w ($cm^2\ s^{-1}$)	1.04×10^{-5}
Cuticular resistance, r_{cl} ($s\ cm^{-1}$)	2.51×10^{4}
Henry's Law constant, H ($Pa\ m^3\ mol^{-1}$)	557

Table 11. Emissions by source category (nearest ton) for each year and inventory scenario.

Source category	Year						
	1990	2000		2010		2020	
		CAA	NON-CAA	CAA	NON-CAA	CAA	NON-CAA
Major	2,495	681	3,075	709	3,262	826	3,775
Area & other	2,861	499	3,394	495	3,366	557	3,846
Nonroad	792	617	962	407	1,199	418	1,433
Onroad	2,371	759	1,537	327	1,448	279	1,971
TOTAL	8,519	2,556	8,968	1,938	9,275	2,080	11,025

Table 12. Percent differences of emissions by source category between each future year and 1990 and between NON-CAA and CAA emissions for each year.

Differences between future year inventories and 1990	Source Category				
	Major	Area & other	Onroad mobile	Nonroad mobile	Total (all sources)
2000 CAA	-73%	-83%	-68%	-22%	-70%
2000 NON-CAA	23%	19%	-35%	21%	5%
2010 CAA	-72%	-83%	-86%	-49%	-77%
2010 NON-CAA	31%	18%	-39%	51%	9%
2020 CAA	-67%	-81%	-88%	-47%	-76%
2020 NON-CAA	51%	34%	-17%	81%	29%
NON-CAA - CAA	**Source Category**				
	Major	**Area & other**	**Onroad mobile**	**Nonroad mobile**	**Total (all sources)**
2000	352%	580%	102%	56%	251%
2010	360%	580%	342%	194%	378%
2020	357%	591%	606%	243%	430%

Table 13. AERMOD control options (CO pathway of runstream file).

Option	Description	Value(s)	Comments
MODELOPT	Modeling options	CONC TOXICS	Calculate concentrations and use TOXICS option
AVERTIME	Concentration average times	1 24 ANNUAL	Calculate hourly, daily, and annual average concentrations
URBANOPT	Urban dispersion options: urban population, descriptor, and surface roughness length (m)	2800000 HOU 0.6	Use urban population of 2,800,000 and surface roughness length of 0.6 m
POLLUTID	Identifier to identify pollutant	BENZENE	Benzene
HALFLIFE	Half life of pollutant; exponential decay in seconds	1209600	Half-life of 1,209,600 seconds
GDSEASON	Seasons for each month to use for deposition calculations	3 3 3 5 5 1 1 1 1 2 3 3	January – March, November, & December: late autumn; April & May: transitional spring; June – September, midsummer; October autumn with unharvested crops
GDLANUSE	Landuse use by sector	36*1	Use urban landuse for all sectors

Table 14. County average concentrations (µg per m^3). County specific background concentrations listed with county names in parentheses.

Year	Scenario	Source category	County (Background)		
			Brazoria (3.63x10^{-1})	Galveston (3.97x10^{-1})	Harris (4.64x10^{-1})
1990	Base	Stationary	8.08x10^{-1}	7.27x10^{0}	6.97x10^{-1}
		Mobile	8.90x10^{-2}	2.21x10^{-1}	4.15x10^{-1}
		Total	1.26x10^{0}	7.89x10^{0}	1.58x10^{0}
2000	CAA	Stationary	1.04x10^{-1}	1.35x10^{-1}	1.37x10^{-1}
		Mobile	3.63x10^{-2}	6.60x10^{-2}	1.82x10^{-1}
		Total	5.04x10^{-1}	5.98x10^{-1}	7.83x10^{-1}
	NON- CAA	Stationary	1.40x10^{0}	1.16x10^{0}	9.10x10^{-1}
		Mobile	6.25x10^{-2}	1.11x10^{-1}	3.46x10^{-1}
		Total	1.83x10^{0}	1.66x10^{0}	1.72x10^{0}
2010	CAA	Stationary	1.15x10^{-1}	1.32x10^{-1}	1.42x10^{-1}
		Mobile	1.73x10^{-2}	3.19x10^{-2}	9.56x10^{-2}
		Total	4.95x10^{-1}	5.62x10^{-1}	7.02x10^{-1}
	NON- CAA	Stationary	1.39x10^{0}	1.17x10^{0}	9.19x10^{-1}
		Mobile	6.31x10^{-2}	1.04x10^{-1}	3.70x10^{-1}
		Total	1.81x10^{0}	1.67x10^{0}	1.75x10^{0}
2020	CAA	Stationary	1.28x10^{-1}	1.48x10^{-1}	1.62x10^{-1}
		Mobile	1.55x10^{-2}	2.82x10^{-2}	9.11x10^{-2}
		Total	5.07x10^{-1}	5.73x10^{-1}	7.17x10^{-1}
	NON- CAA	Stationary	1.58x10^{0}	1.34x10^{0}	1.05x10^{0}
		Mobile	7.79x10^{-2}	1.24x10^{-1}	4.71x10^{-1}
		Total	2.02x10^{0}	1.86x10^{0}	1.99x10^{0}

Table 15. Percent differences for domain average AERMOD concentrations by source category between each future year and 1990 and between NON-CAA and CAA concentrations for each year.

Differences between future year inventories and 1990	Source Category					
	Major	Area & other	Onroad mobile	Nonroad mobile	Background	Total (all sources)
2000 CAA	-60%	-94%	-68%	-20%	0%	-65%
2000 NON-CAA	44%	-38%	-33%	30%	0%	-20%
2010 CAA	-58%	-94%	-86%	-49%	0%	-69%
2010 NON-CAA	53%	-38%	-37%	65%	0%	-19%
2020 CAA	-52%	-94%	-89%	-45%	0%	-68%
2020 NON-CAA	76%	-30%	-17%	99%	0%	-8%
NON-CAA - CAA	**Source Category**					
	Major	**Area & other**	**Onroad mobile**	**Nonroad mobile**	**Background**	**Total (all sources)**
2000	259%	1006%	109%	63%	0%	131%
2010	264%	961%	355%	222%	0%	159%
2020	265%	984%	645%	260%	0%	187%

Table 16. County average HAPEM6 exposure concentrations (µg per m^3).

Year	Scenario	Source category	County		
			Brazoria	**Galveston**	**Harris**
1990	Base	Stationary	7.32x10^{-1}	6.32x10^{0}	5.67x10^{-1}
		Mobile	1.24x10^{-1}	2.29x10^{-1}	4.71x10^{-1}
		Background	2.93x10^{-1}	3.05x10^{-1}	3.74x10^{-1}
		Total	1.15x10^{0}	6.85x10^{0}	1.41x10^{0}
2000	CAA	Stationary	7.52x10^{-2}	1.03x10^{-1}	1.14x10^{-1}
		Mobile	5.05x10^{-2}	7.59x10^{-2}	2.04x10^{-1}
		Background	2.97x10^{-1}	3.31x10^{-1}	3.77x10^{-1}
		Total	4.23x10^{-1}	5.10x10^{-1}	6.95x10^{-1}
	NON-CAA	Stationary	9.74x10^{-1}	7.69x10^{-1}	7.86x10^{-1}
		Mobile	7.71x10^{-2}	1.21x10^{-1}	3.71x10^{-1}
		Background	2.54x10^{-1}	3.09x10^{-1}	3.53x10^{-1}
		Total	1.31x10^{0}	1.20x10^{0}	1.51x10^{0}
2010	CAA	Stationary	8.05x10^{-2}	1.02x10^{-1}	1.18x10^{-1}
		Mobile	2.45x10^{-2}	3.66x10^{-2}	1.05x10^{-1}
		Background	3.01x10^{-1}	3.34x10^{-1}	3.84x10^{-1}
		Total	4.06x10^{-1}	4.73x10^{-1}	6.07x10^{-1}
	NON-CAA	Stationary	9.62x10^{-1}	7.79x10^{-1}	7.88x10^{-1}
		Mobile	7.86x10^{-2}	1.16x10^{-1}	3.86x10^{-1}
		Background	2.55x10^{-1}	3.10x10^{-1}	3.53x10^{-1}
		Total	1.30x10^{0}	1.20x10^{0}	1.53x10^{0}
2020	CAA	Stationary	8.89x10^{-2}	1.13x10^{-1}	1.35x10^{-1}
		Mobile	2.27x10^{-2}	3.31x10^{-2}	1.01x10^{-1}
		Background	3.00x10^{-1}	3.34x10^{-1}	3.83x10^{-1}
		Total	4.12x10^{-1}	4.80x10^{-1}	6.19x10^{-1}
	NON-CAA	Stationary	1.09x10^{0}	8.82x10^{-1}	8.97x10^{-1}
		Mobile	9.78x10^{-2}	1.38x10^{-1}	4.97x10^{-1}
		Background	2.53x10^{-1}	3.08x10^{-1}	3.51x10^{-1}
		Total	1.44x10^{0}	1.33x10^{0}	1.75x10^{0}

Table 17. Percent differences for domain average HAPEM6 concentrations by source category between each future year and 1990 and between NON-CAA and CAA concentrations for each year.

Differences between future year inventories and 1990	**Source Category**					
	Major	**Area & other**	**Onroad mobile**	**Nonroad mobile**	**Background**	**Total (all sources)**
2000 CAA	-57%	-94%	-68%	-2%	1%	-64%
2000 NON-CAA	31%	-32%	-37%	50%	-6%	-20%
2010 CAA	-55%	-94%	-86%	-37%	3%	-68%
2010 NON-CAA	40%	-33%	-41%	91%	-6%	-19%
2020 CAA	-48%	-93%	-88%	-31%	3%	-68%
2020 NON-CAA	61%	-24%	-21%	131%	-6%	-8%
NON-CAA - CAA	**Source Category**					
	Major	**Area & other**	**Onroad mobile**	**Nonroad mobile**	**Background**	**Total (all sources)**
2000	203%	1059%	96%	54%	-7%	122%
2010	212%	995%	318%	204%	-8%	155%
2020	211%	1015%	567%	234%	-9%	185%

Table 18. Average ratio of HAPEM6 to AERMOD concentrations.

Inventory	**Ratios**					
	Major	**Area & other**	**Onroad mobile**	**Nonroad mobile**	**Background**	**Total (all sources)**
1990	0.89	0.89	1.37	0.93	0.80	0.93
2000 CAA	0.91	0.87	1.36	1.13	0.81	0.89
2000 NON-CAA	0.82	1.06	1.25	1.05	0.76	0.91
2010 CAA	0.91	0.88	1.35	1.13	0.83	0.87
2010 NON-CAA	0.82	1.05	1.22	1.06	0.76	0.90
2020 CAA	0.92	0.88	1.38	1.17	0.83	0.87
2020 NON-CAA	0.81	1.04	1.22	1.07	0.75	0.91

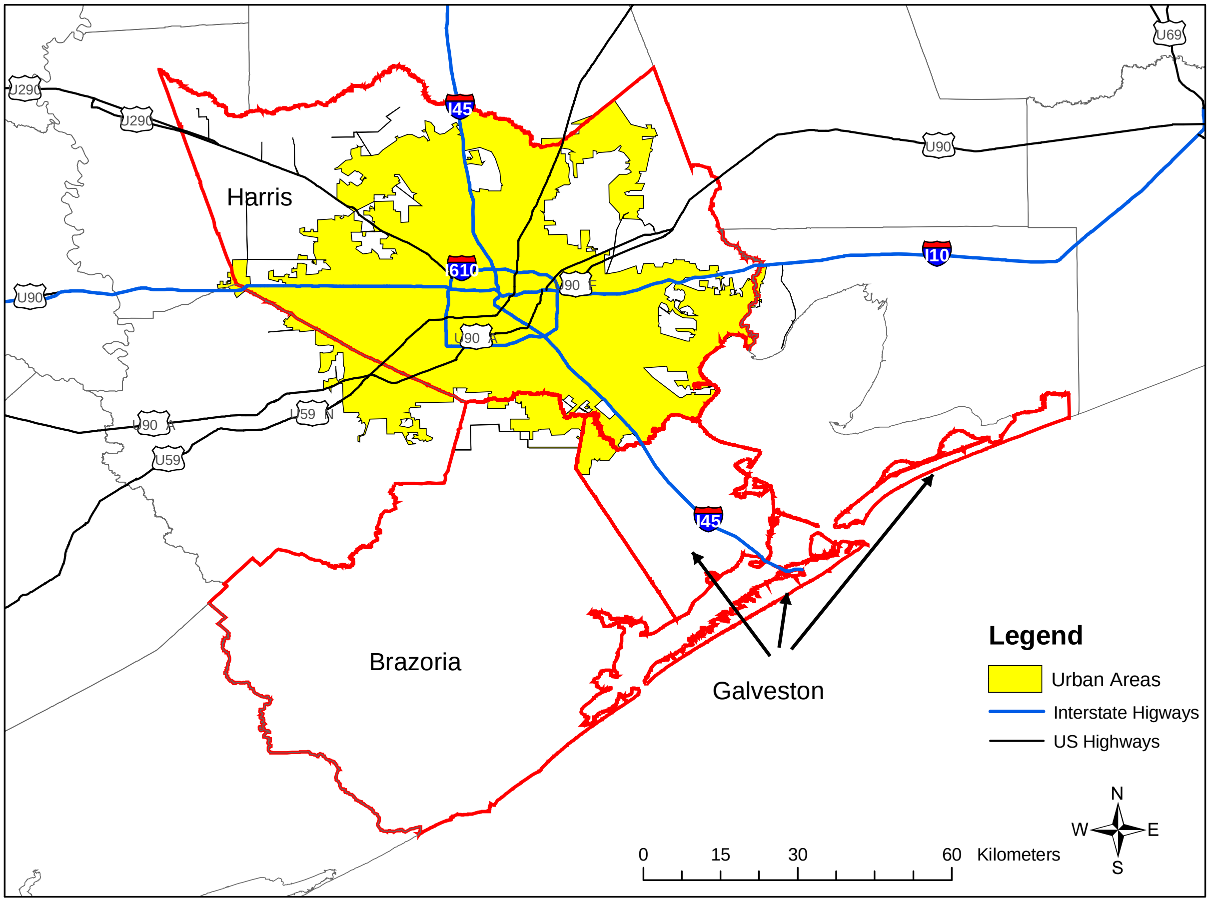


Figure 1. Domain of modeling study. The three counties of emphasis are labeled and outlined in red and the Houston metropolitan area is shown in yellow. Key roads are also shown.

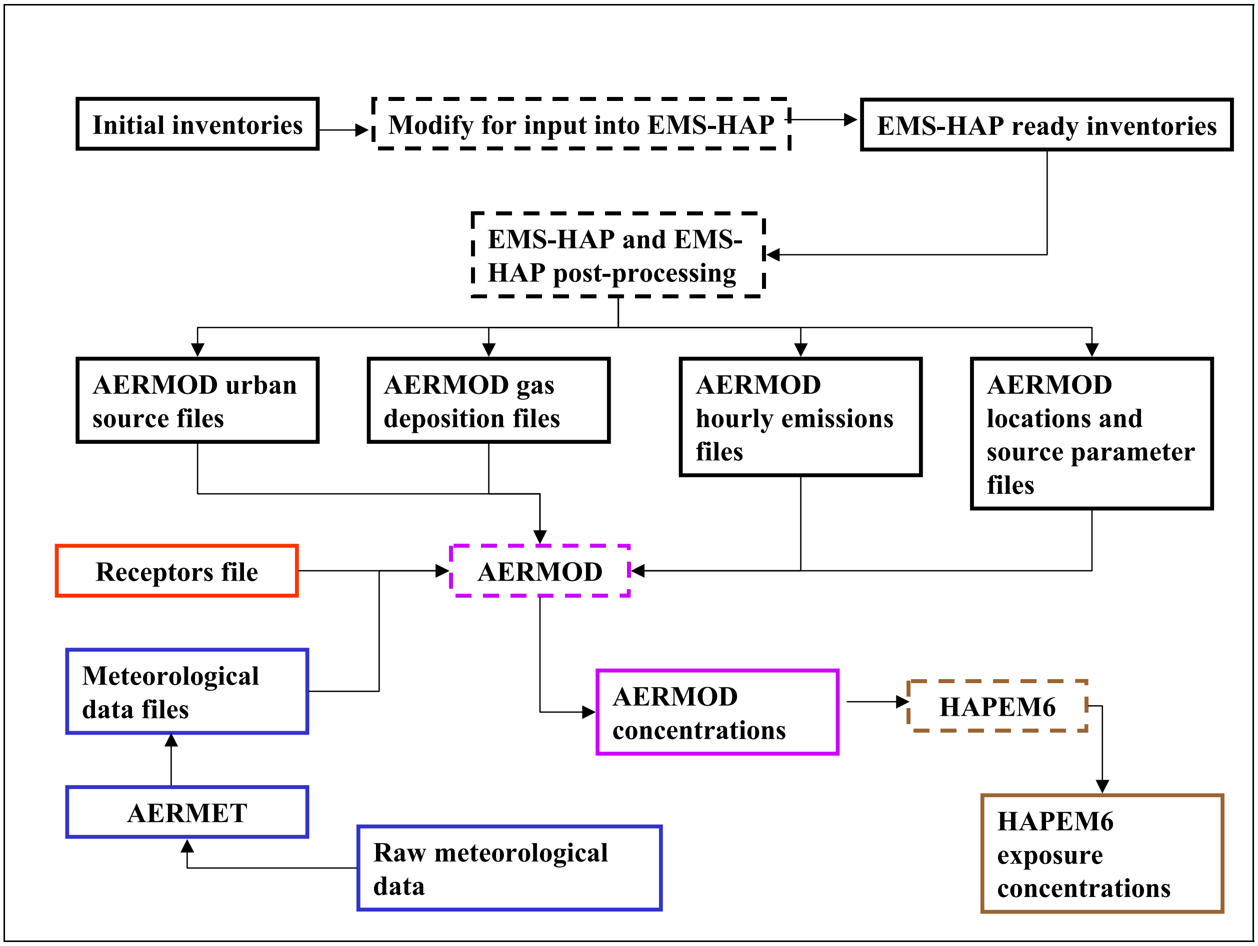


Figure 2. General overview of processing steps from emissions to HAPEM6 exposure concentrations.

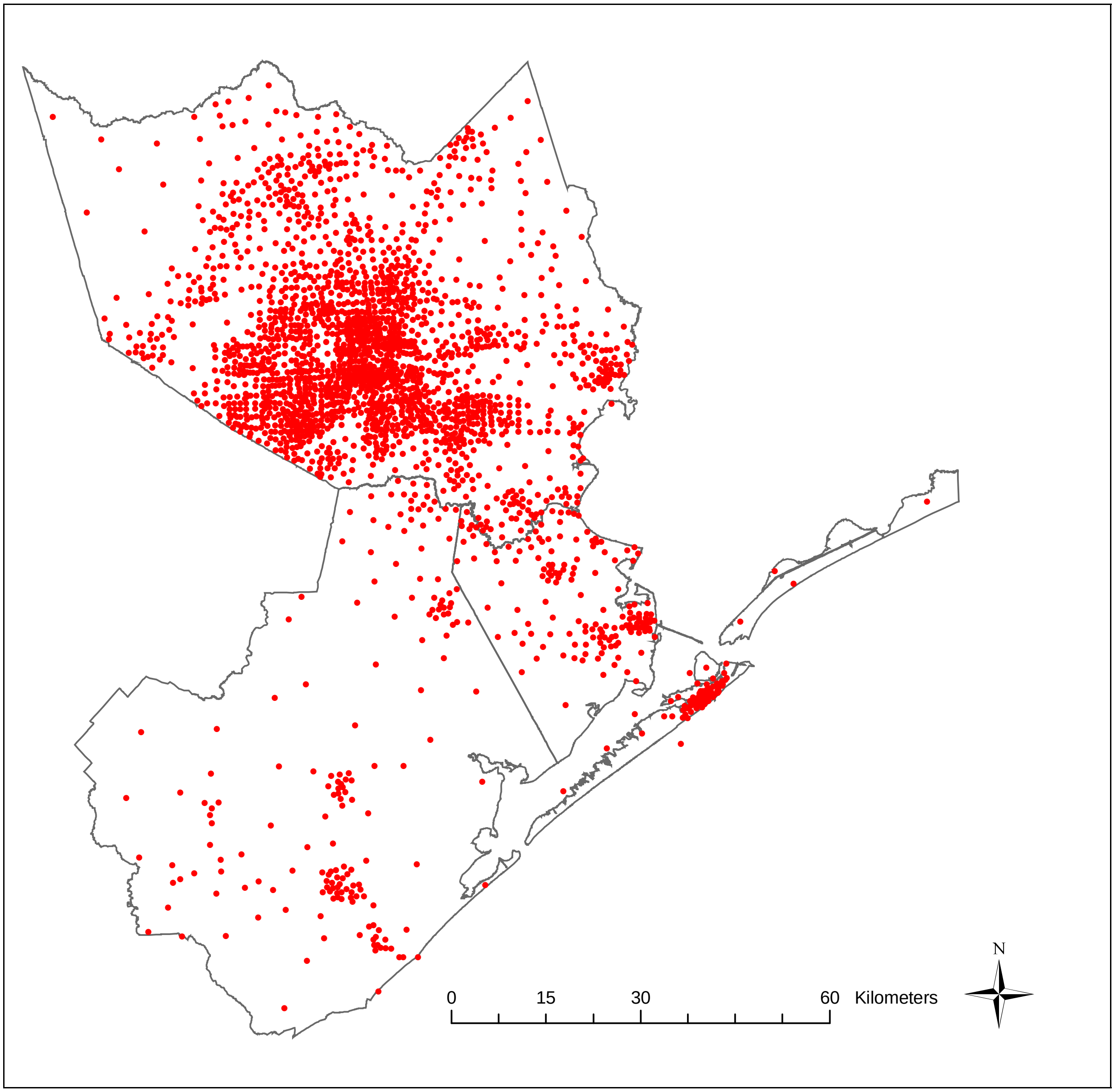


Figure 3. Receptors for 1990 simulation based on 1990 census block group centroids.

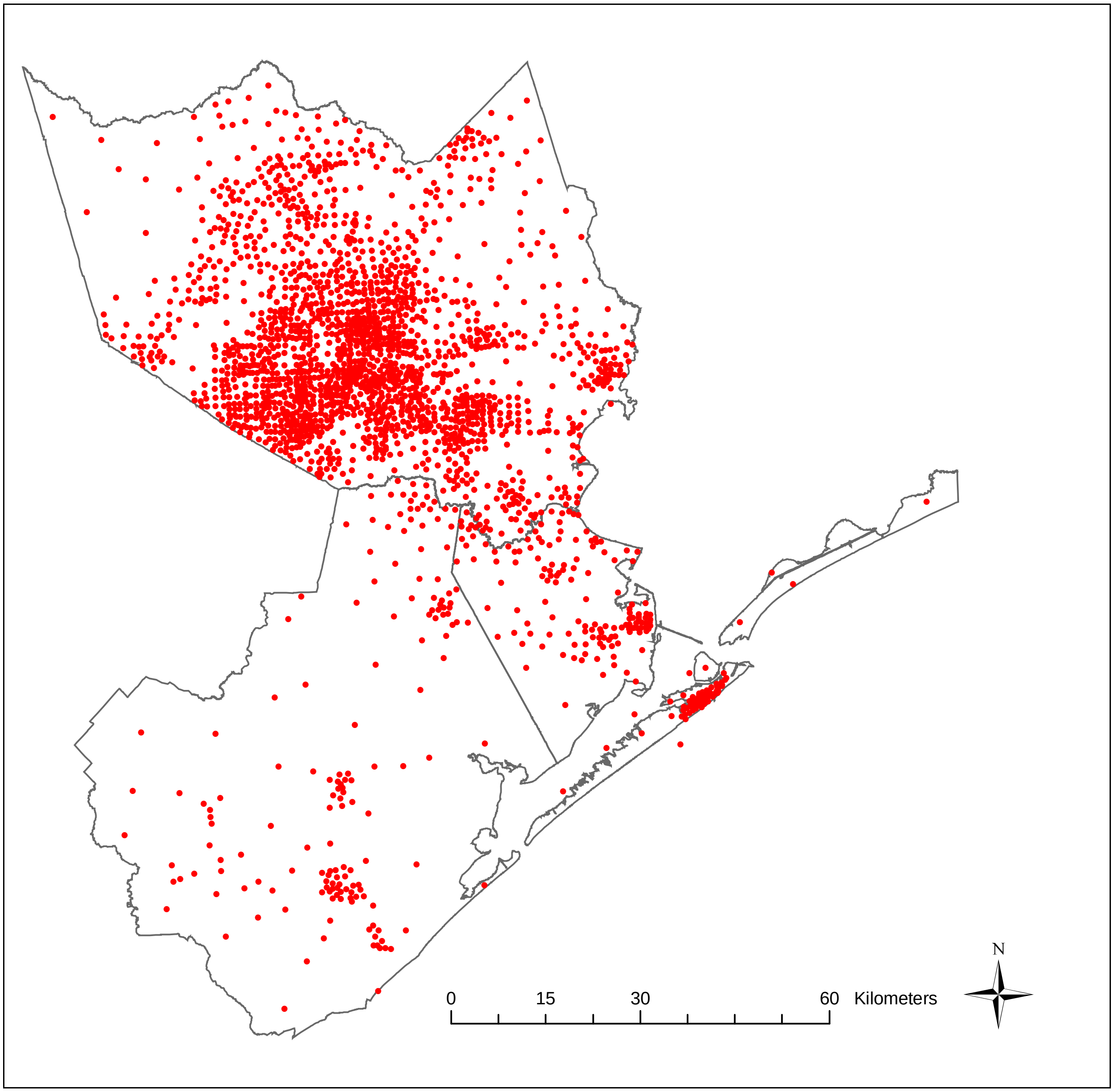


Figure 4. Receptors for simulations of the years 2000, 2010, and 2020 based on 2000 census block group centroids.

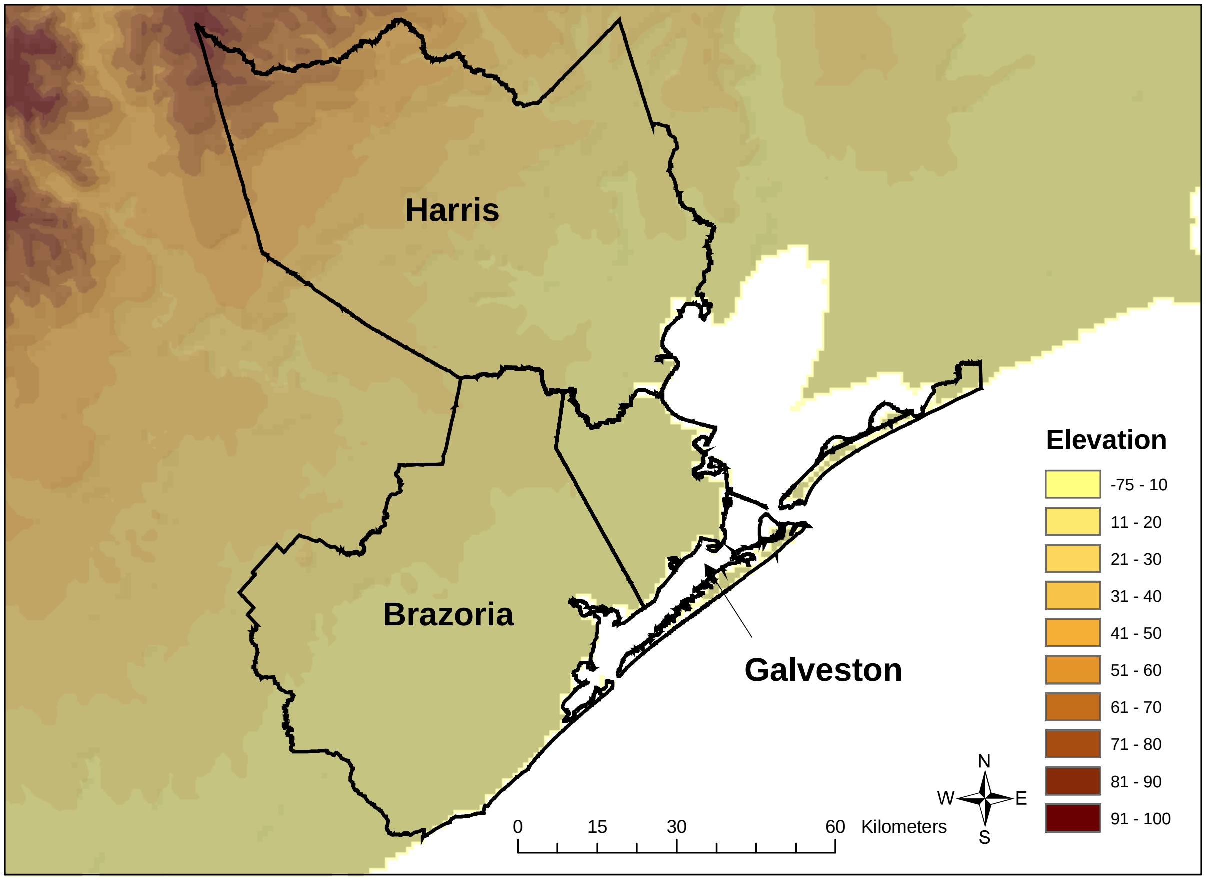


Figure 5. Elevation (m) of the Houston domain.

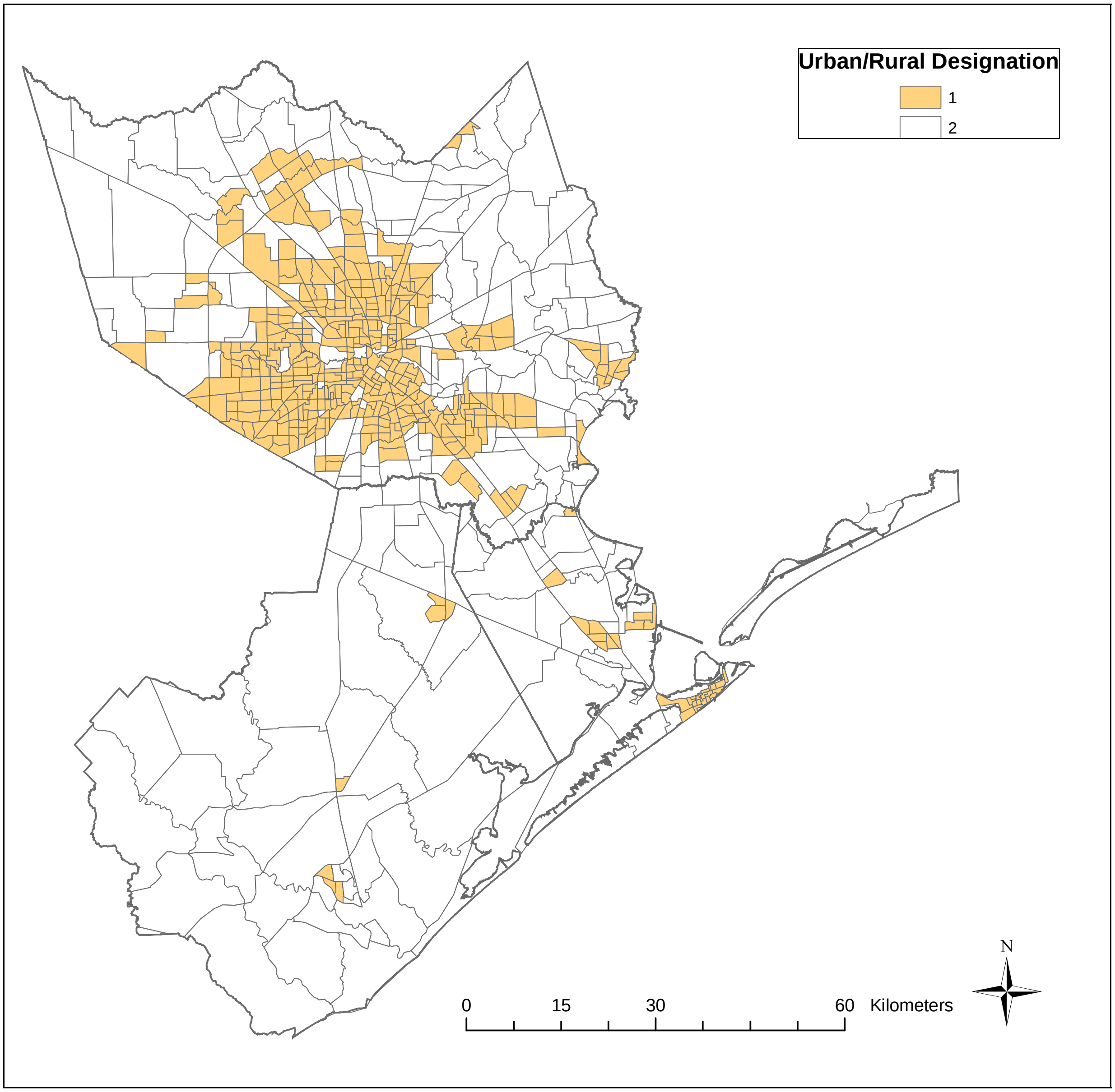


Figure 6. 1990 urban (brown) and rural (white) census tracts.

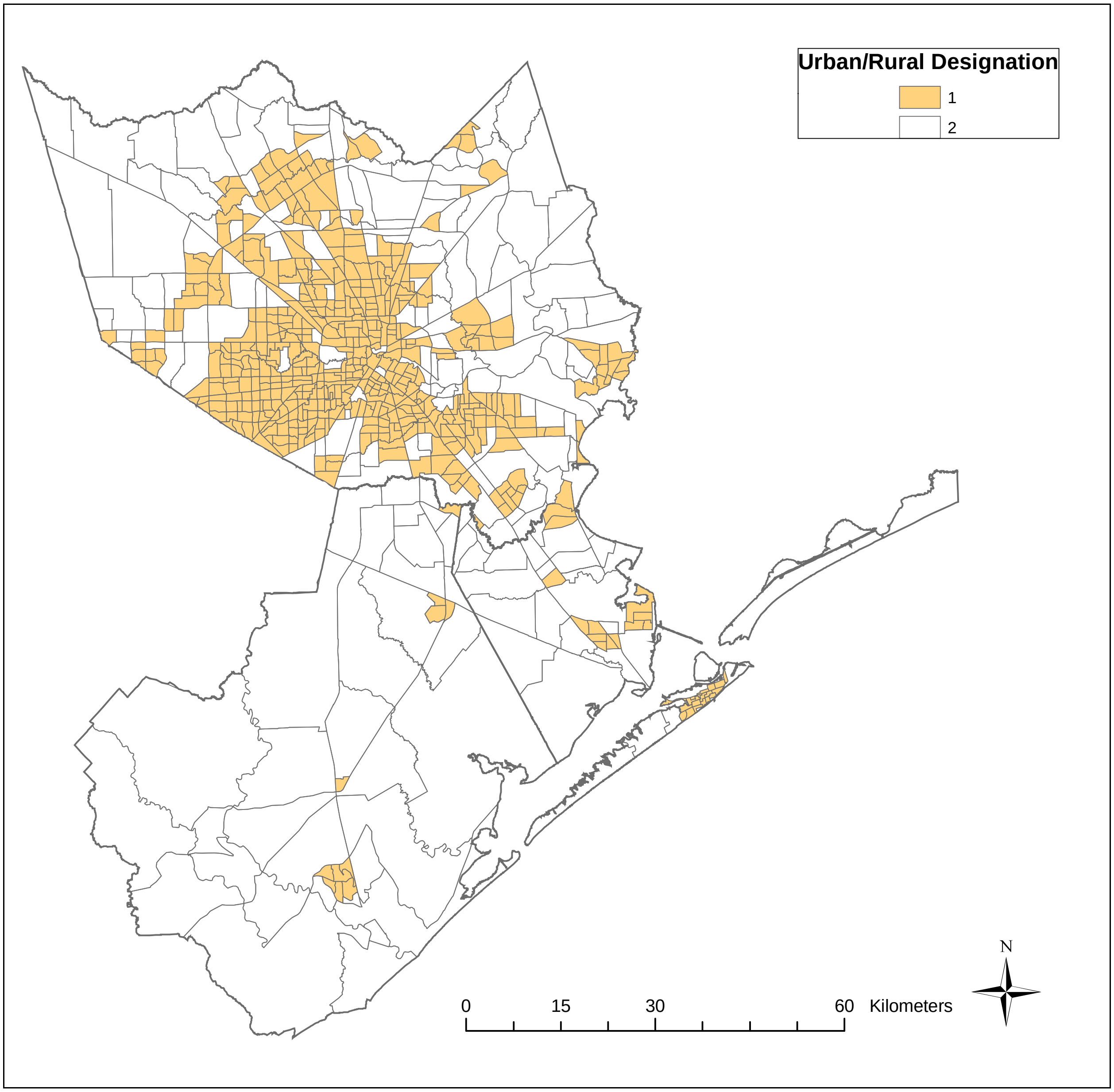


Figure 7. Year 2000 urban (brown) and rural (white) census tracts.

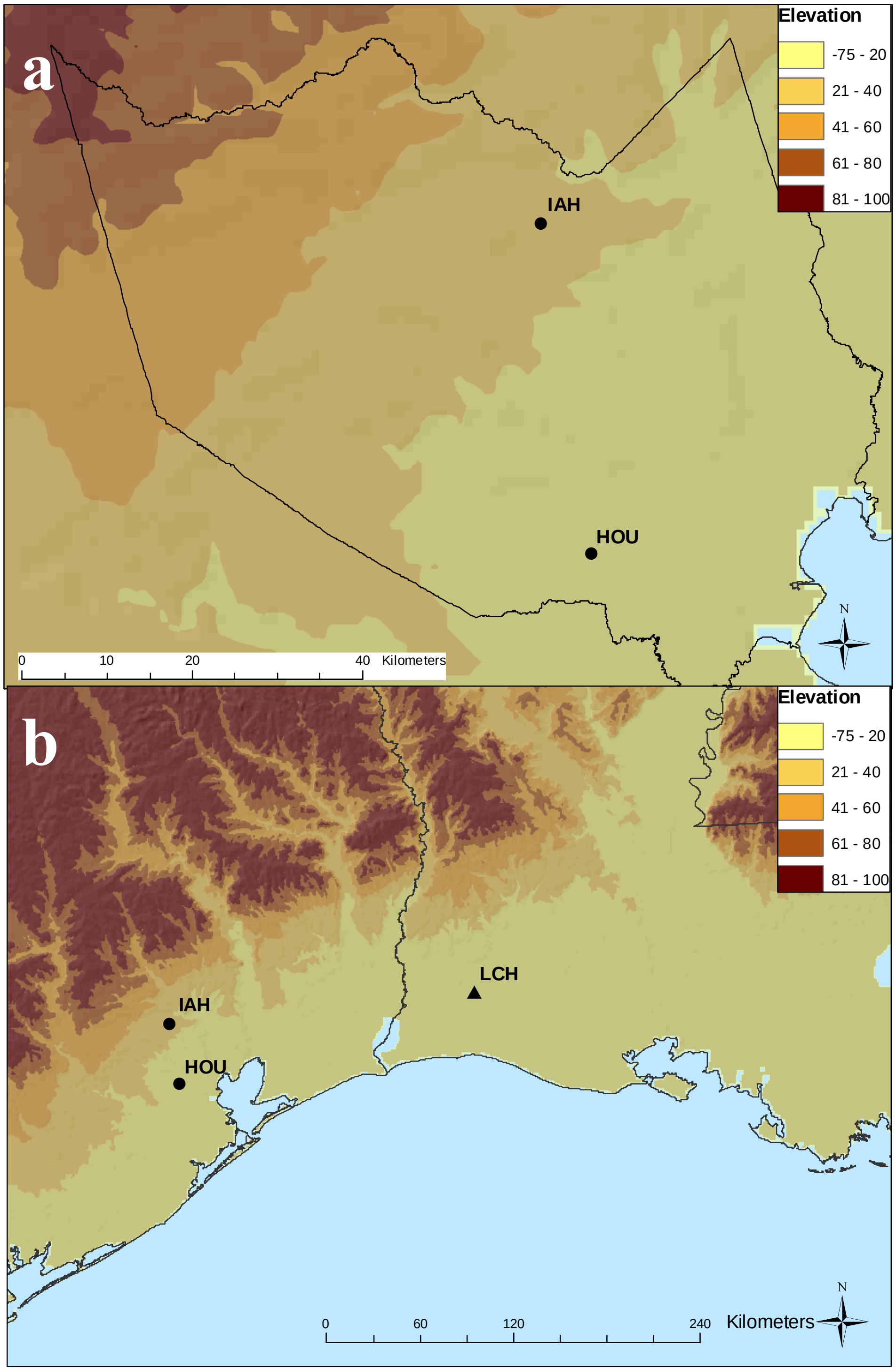


Figure 8. Locations of a) surface meteorological stations for 1990 and 2000 and b) relative locations of surface and upper air stations.

(IAH = George Bush Intercontinental Airport; HOU = Hobby Field; and LCH = Lake Charles)

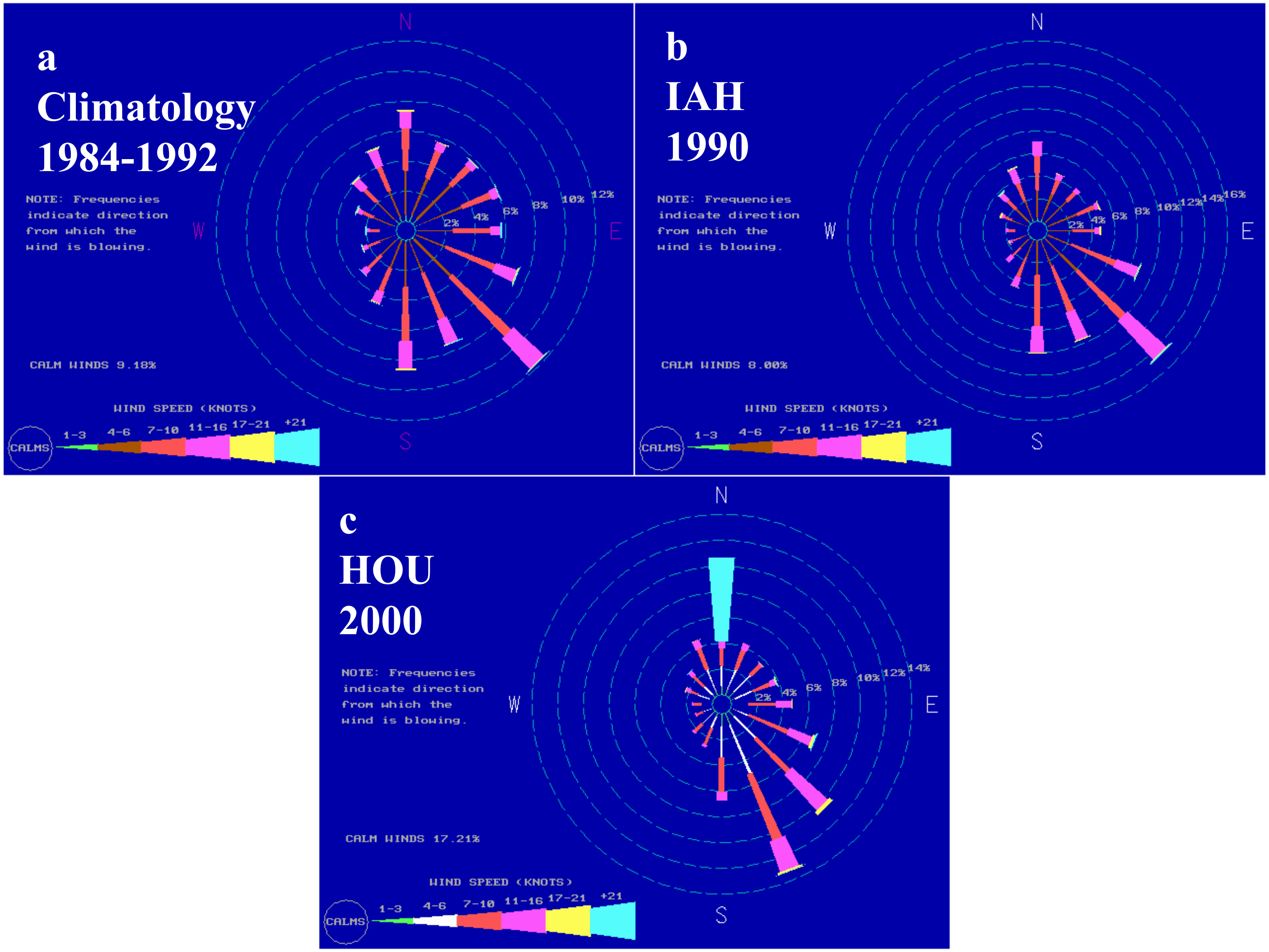
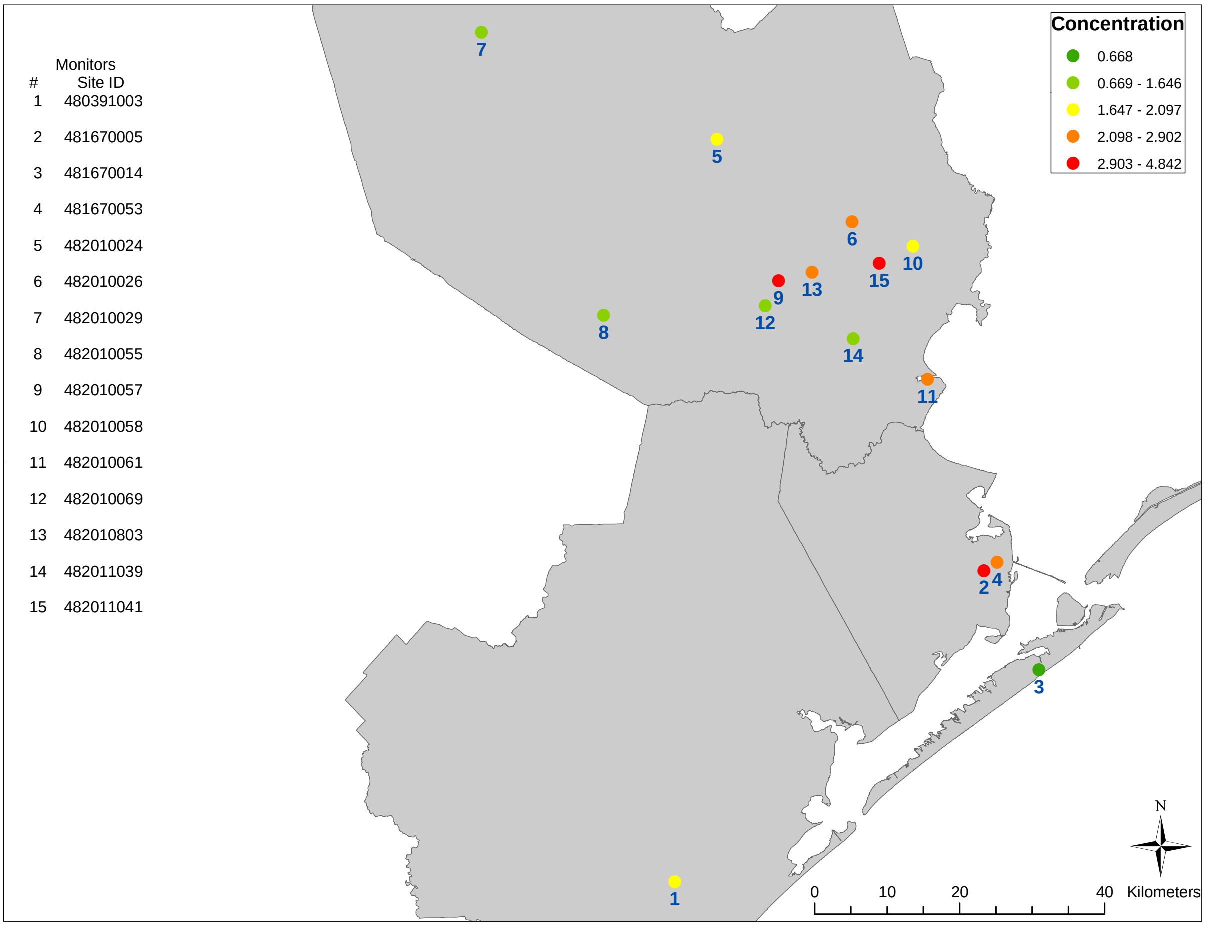


Figure 9. Wind roses for a) 8 year climatology for IAH, b) 1990 IAH, and c) HOU 2000. Winds are generally from the southeast.



Monitors

#	Site ID
1	480391003
2	481670005
3	481670014
4	481670053
5	482010024
6	482010026
7	482010029
8	482010055
9	482010057
10	482010058
11	482010061
12	482010069
13	482010803
14	482011039
15	482011041

Figure 10. Monitor locations and annual average concentrations (µg per m^3) for the year 2000.

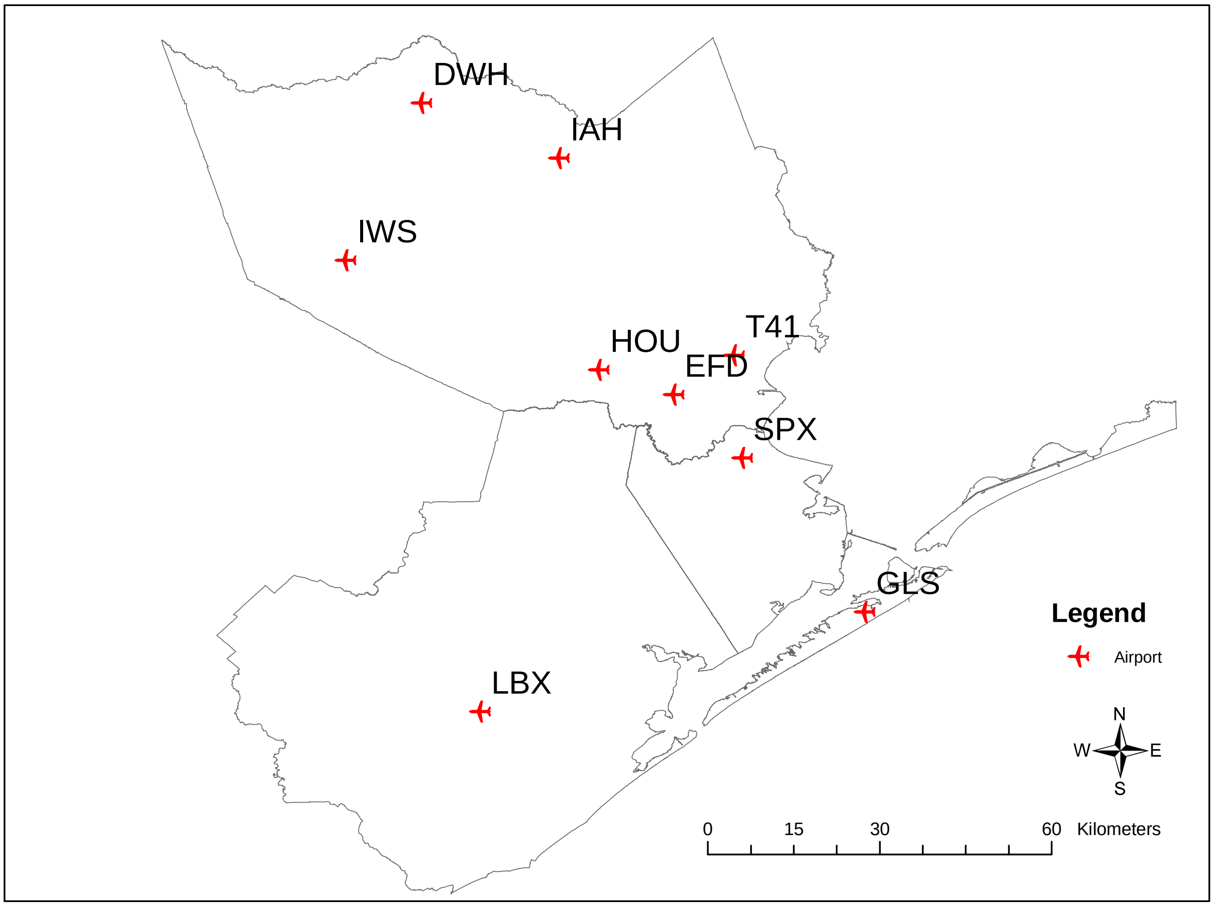


Figure 11. Locations of airports in Houston domain.

FIPS	Airport	Three letter identifier
48039	Brazoria County Airport	LBX
48167	Scholes International Airport at Galveston	GLS
48167	Houston Gulf	SPX
48201	David Wayne Hooks Memorial Airport	DWH
48201	Ellington Field Airport	EFD
48201	William P Hobby Airport	HOU
48201	George Bush Intercontinental Airport	IAH
48201	West Houston Airport	IWS
48201	La Porte Municipal Airport	T41

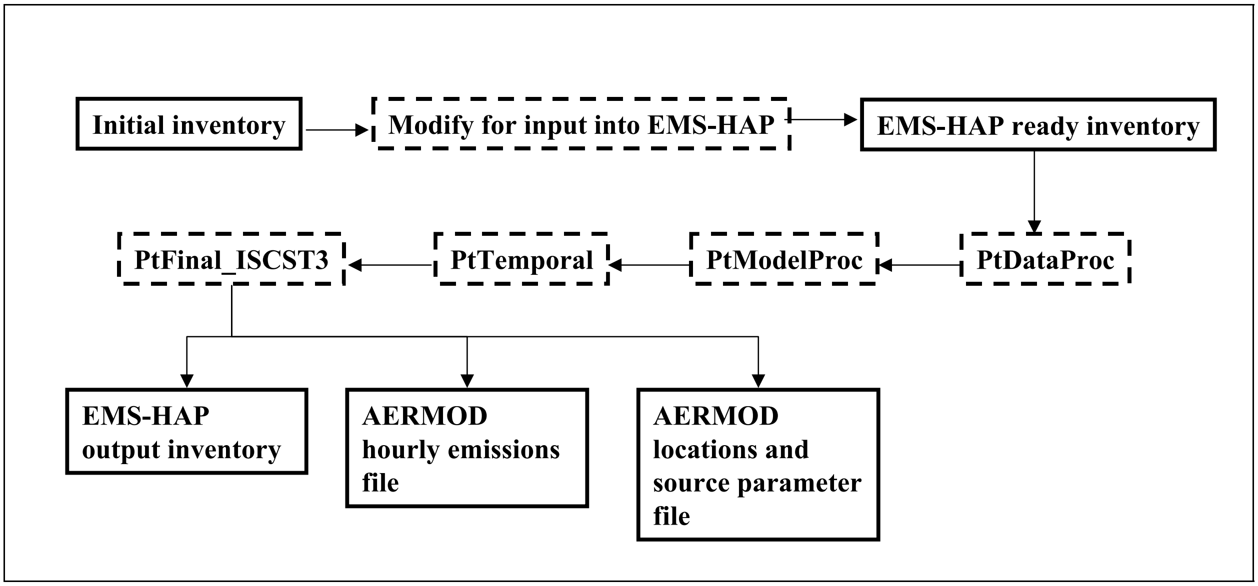


Figure 12. Point inventory EMS-HAP processing.

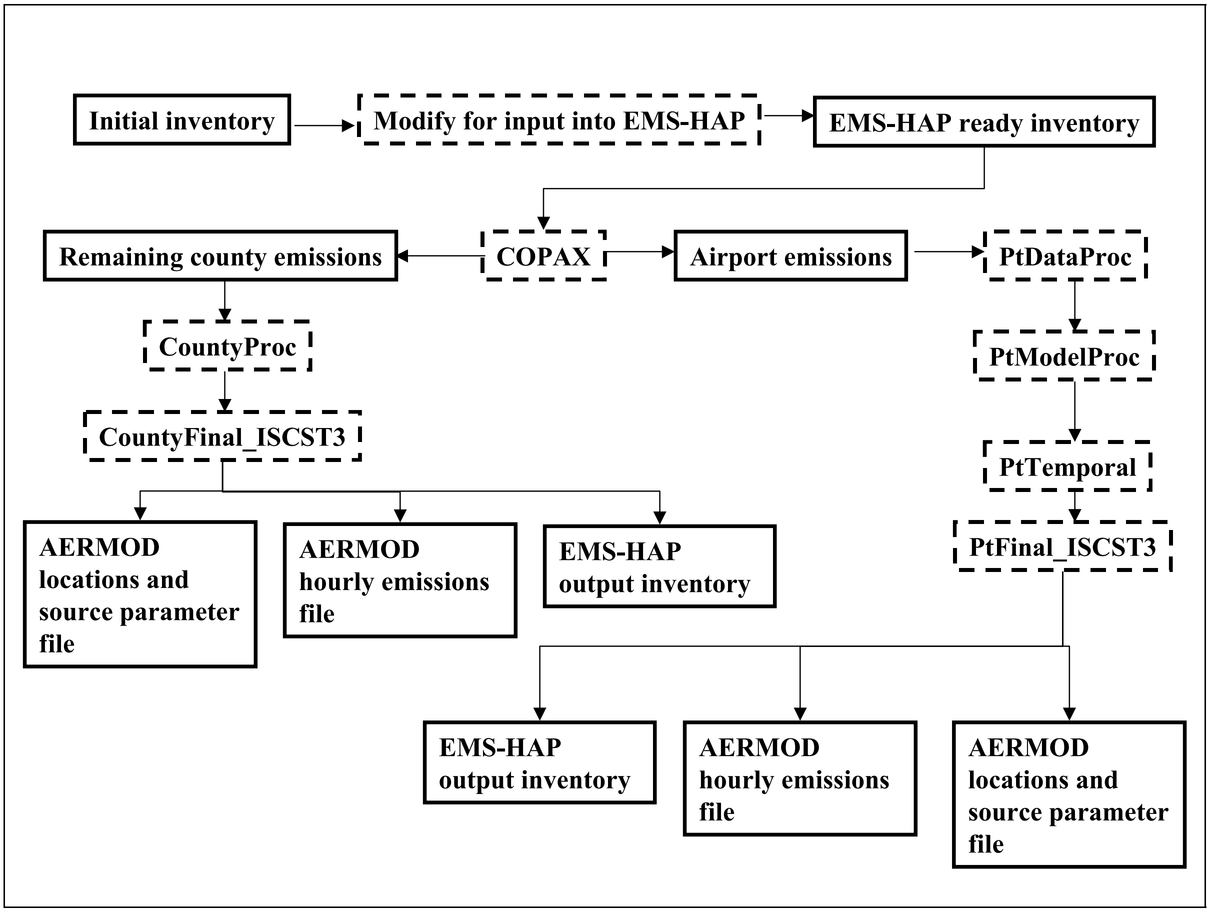


Figure 13. Emissions processing of nonpoint and nonroad inventories.

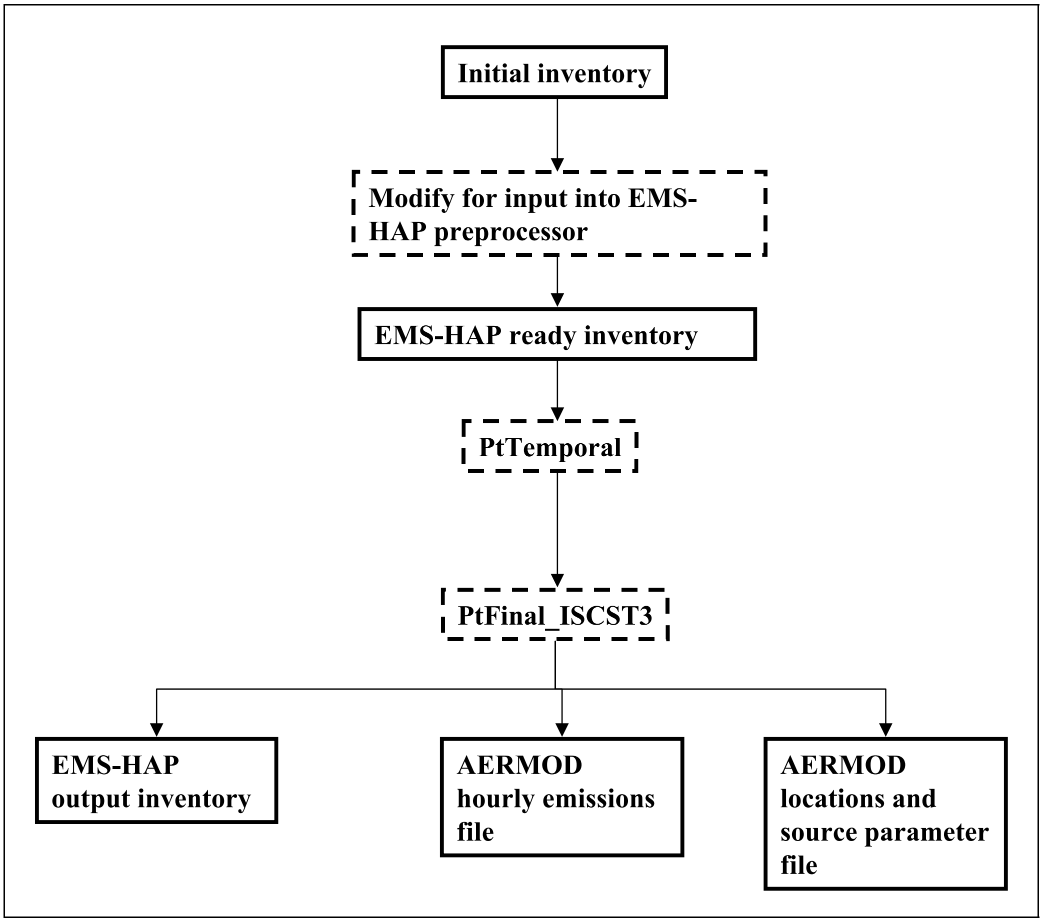


Figure 14. Onroad inventory emissions processing.

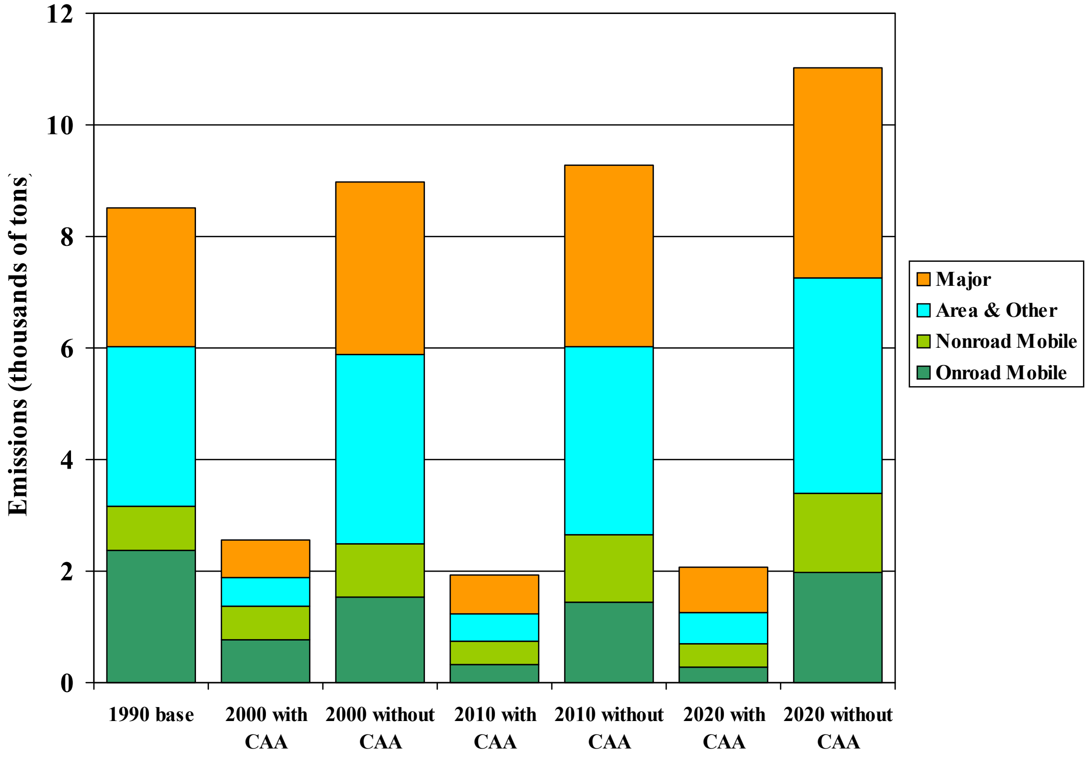


Figure 15. Major, area & other, onroad, and nonroad emissions (tons) for each year and inventory type.

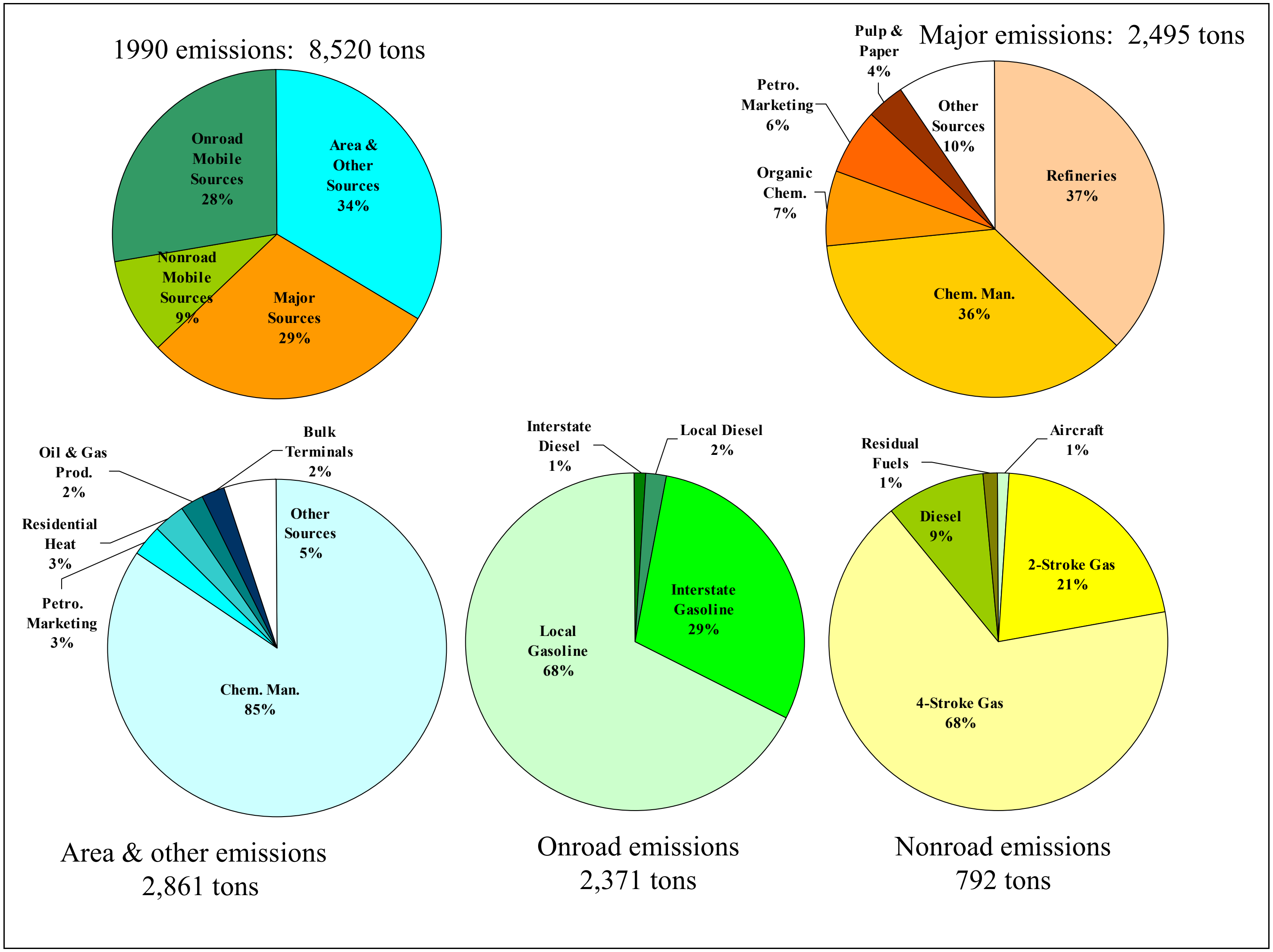


Figure 16. Breakdown of 1990 emissions by source category and within each source category. "Chemical Manufacturing" is included in the area source inventory. With-CAA emissions include Chemical Manufacturing in the point (major source) inventory.

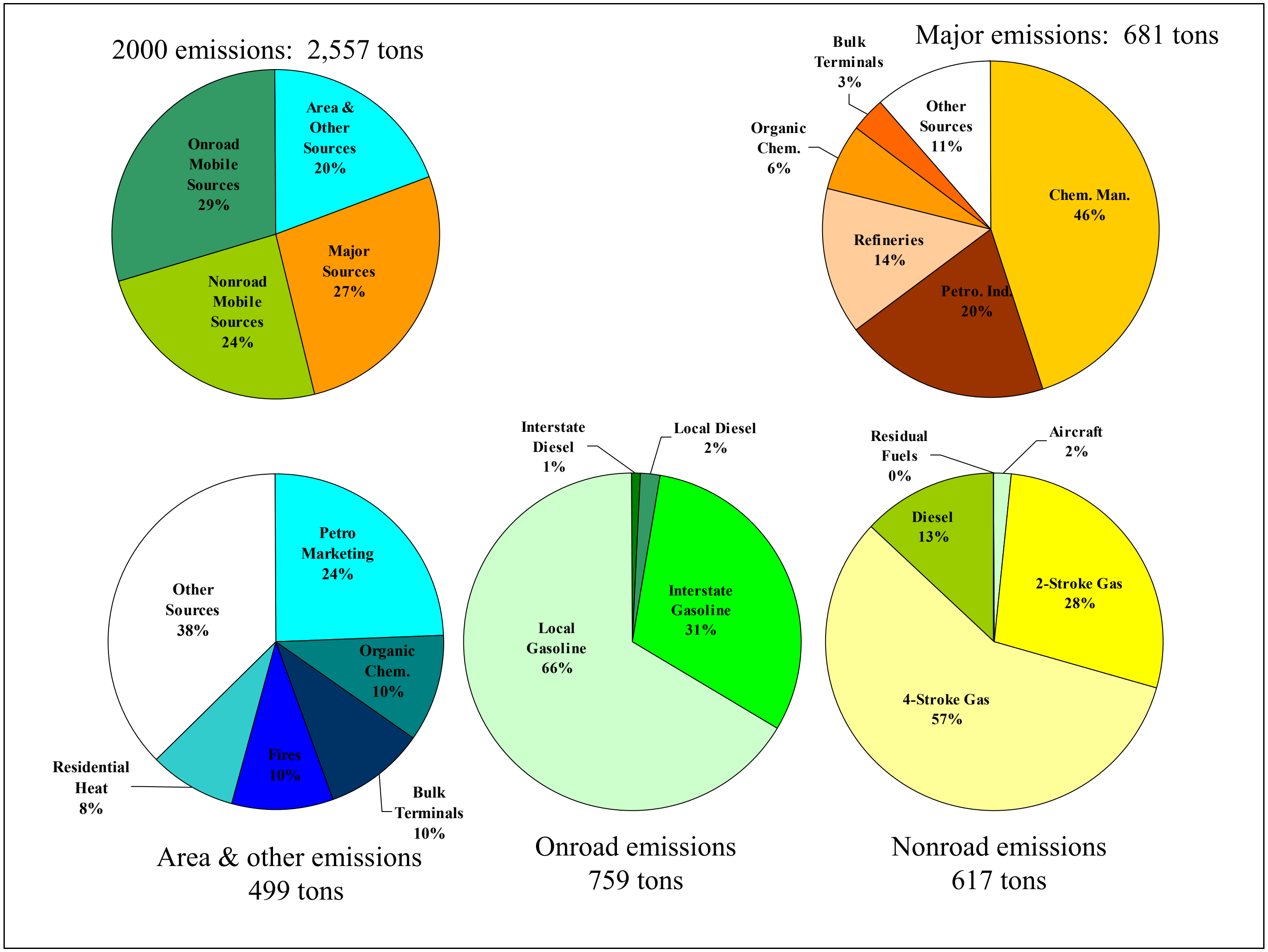


Figure 17. Breakdown of 2000 CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

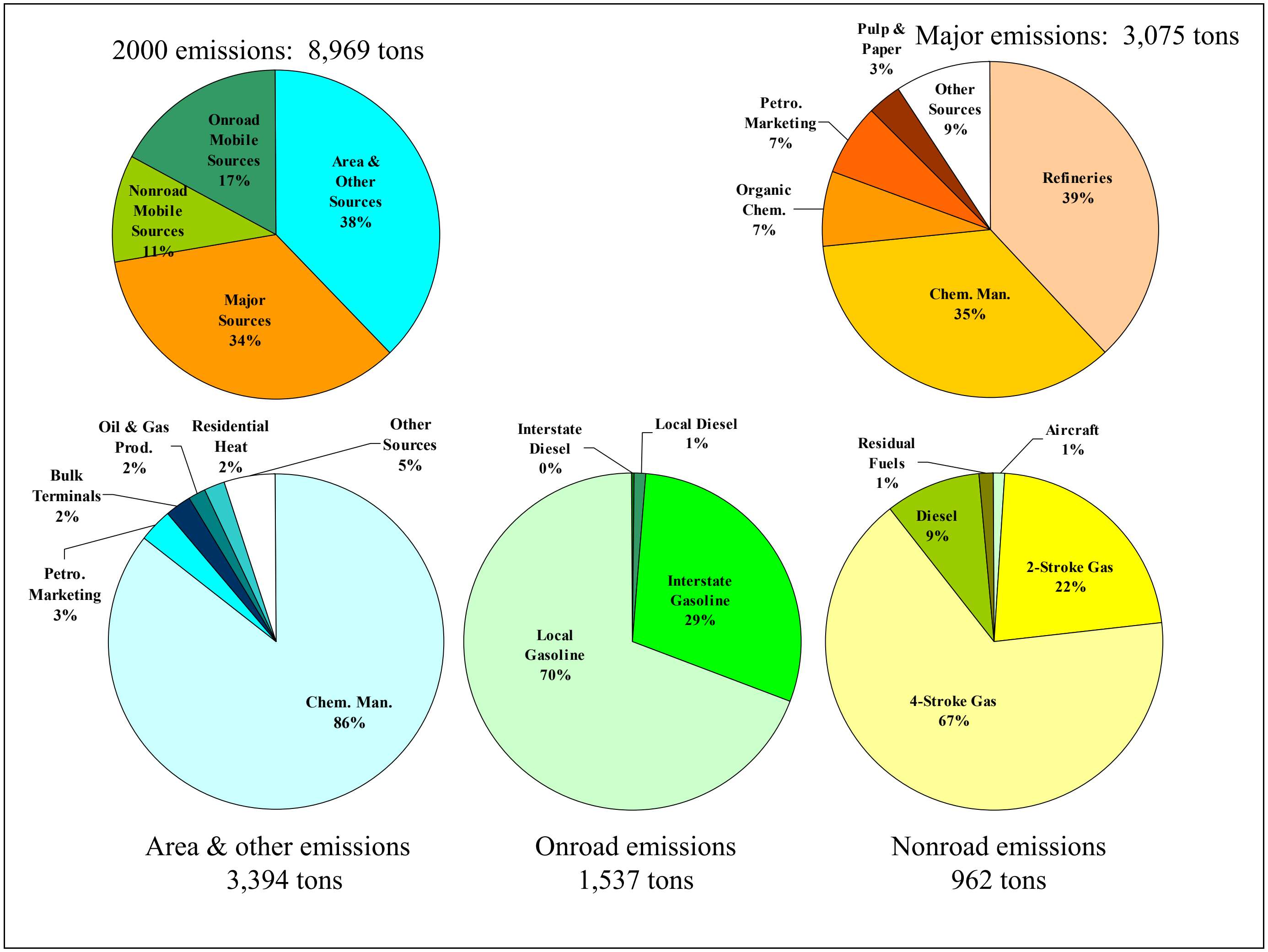


Figure 18. Breakdown of 2000 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

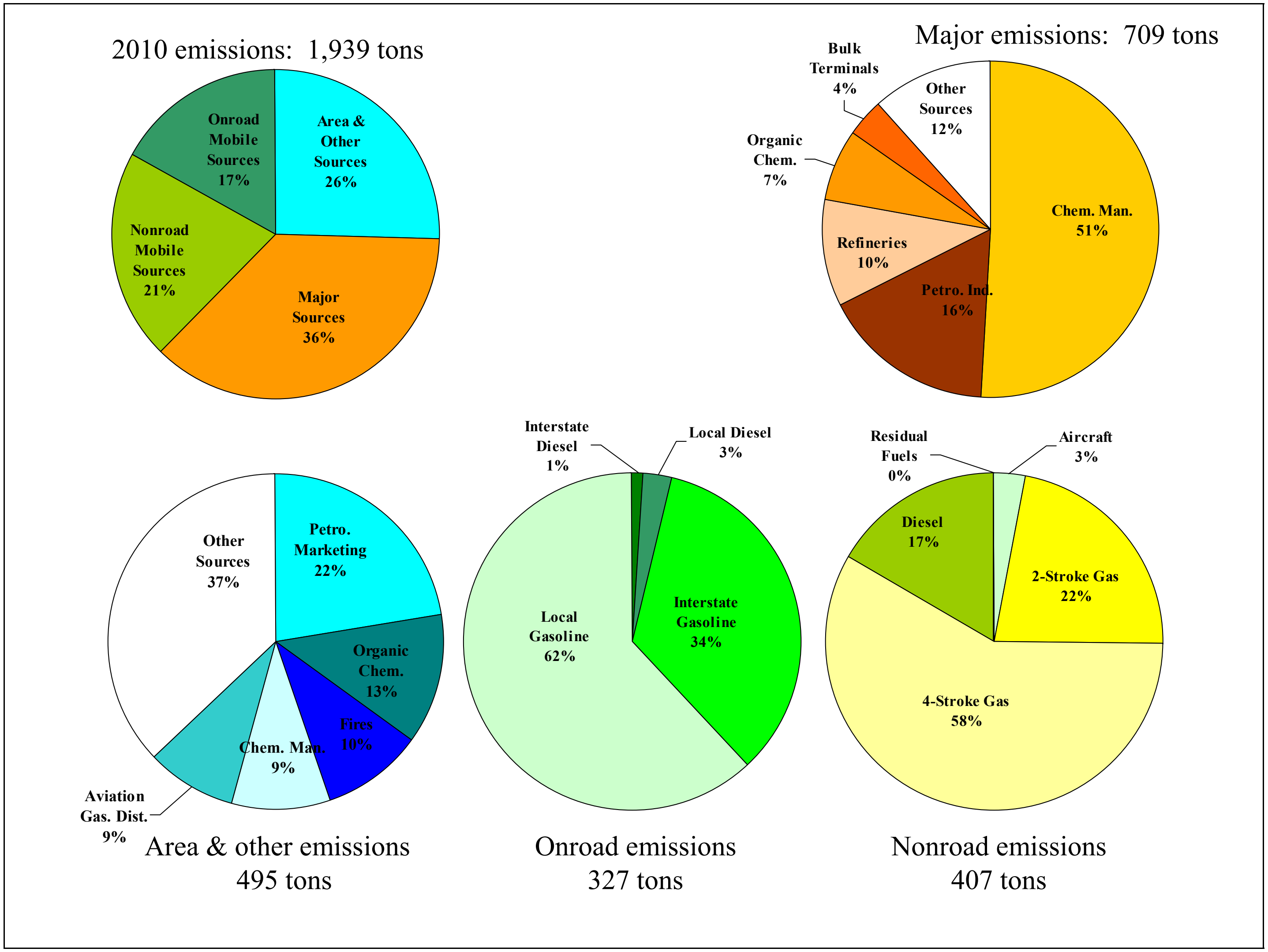


Figure 19. Breakdown of 2010 CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

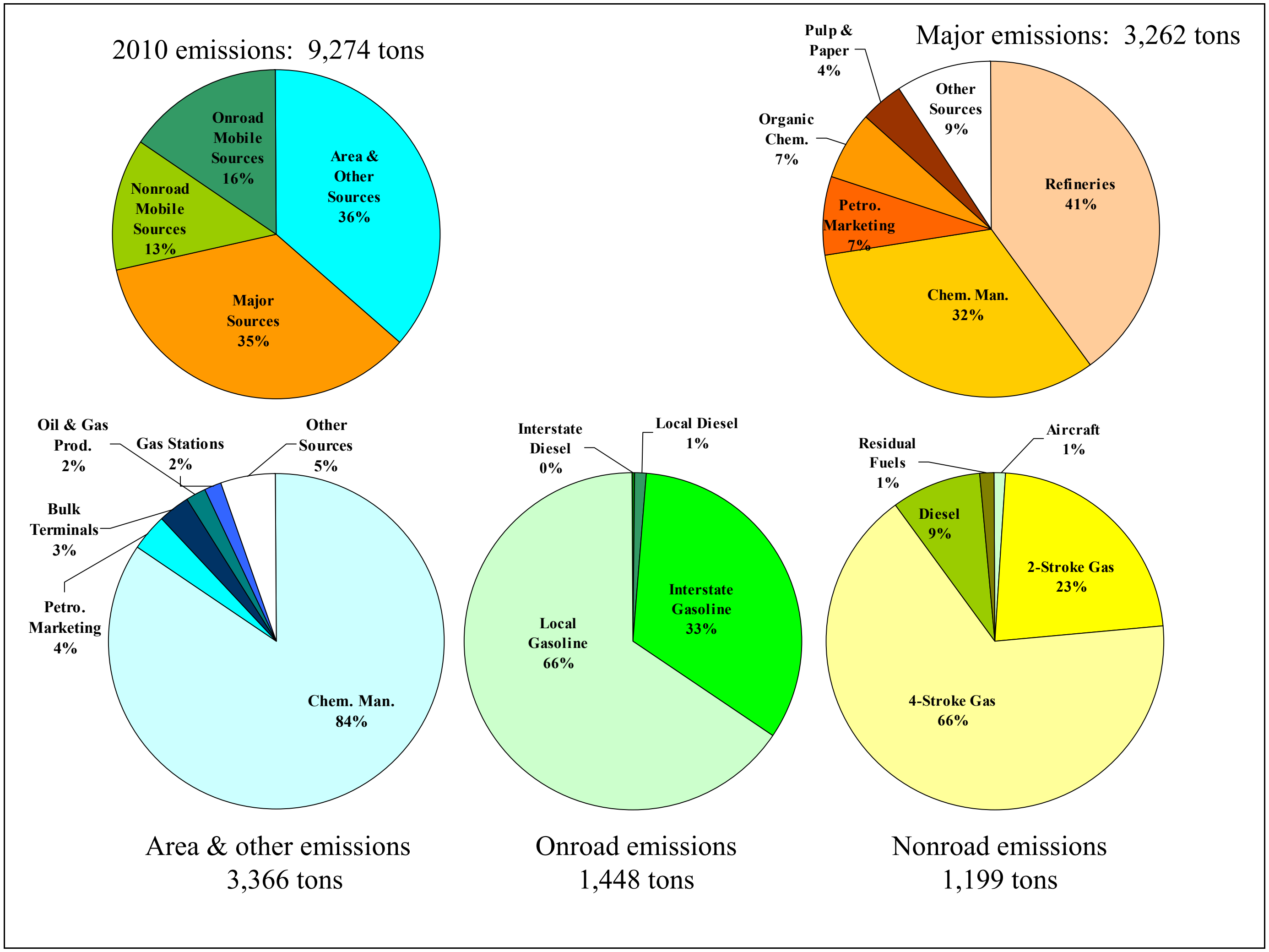


Figure 20. Breakdown of 2010 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

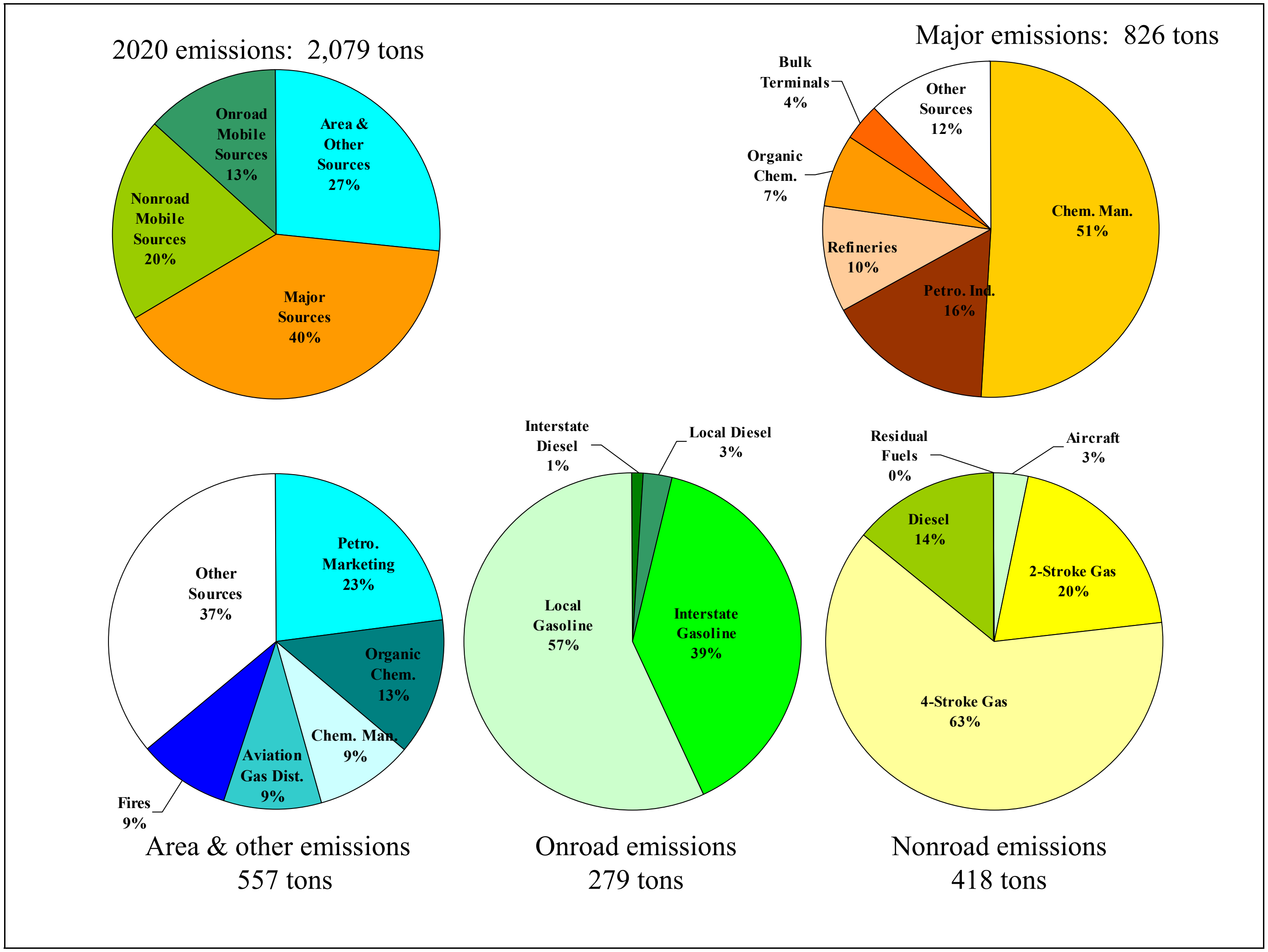


Figure 21. Breakdown of 2020 CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

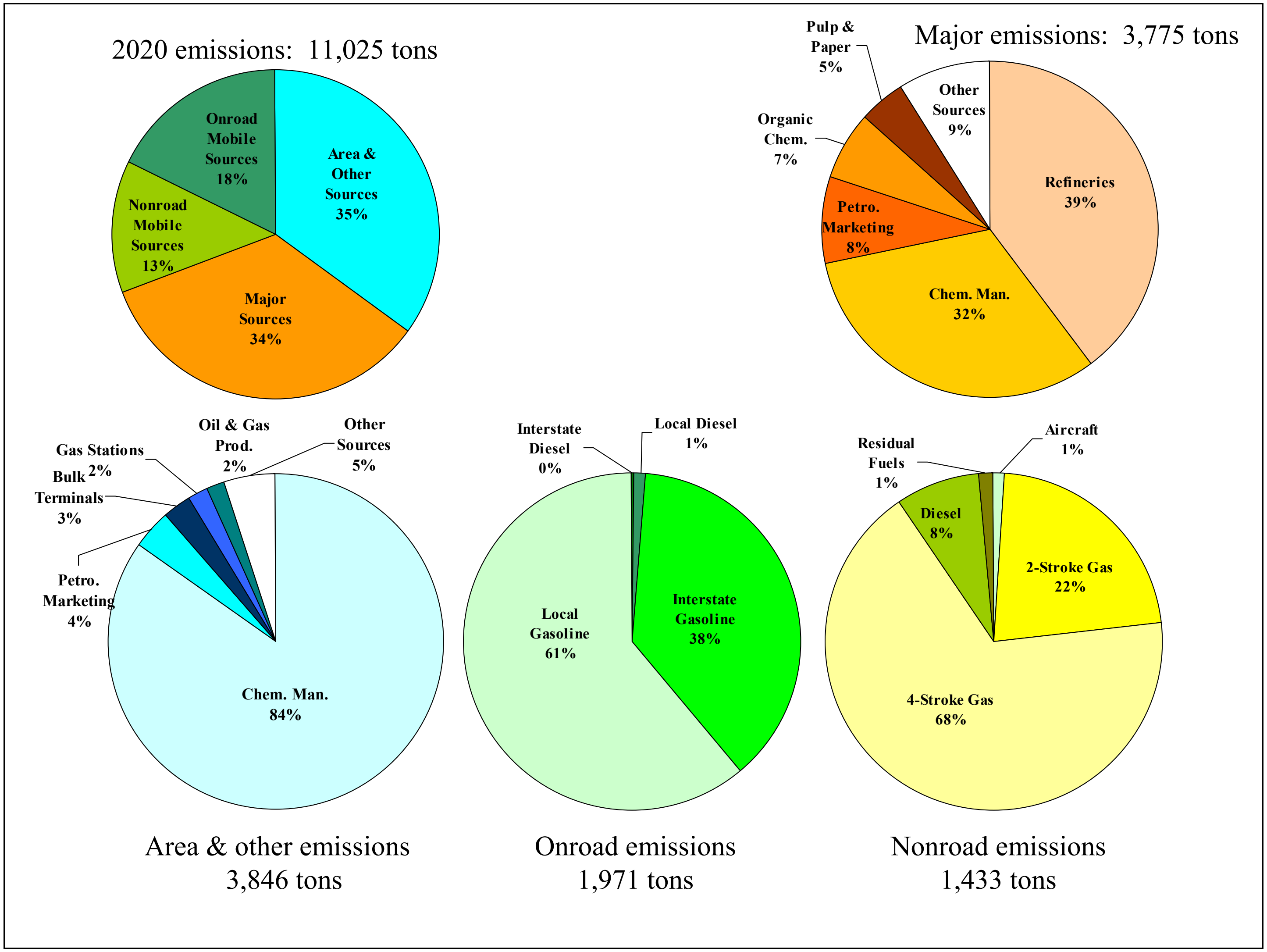


Figure 22. Breakdown of 2020 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

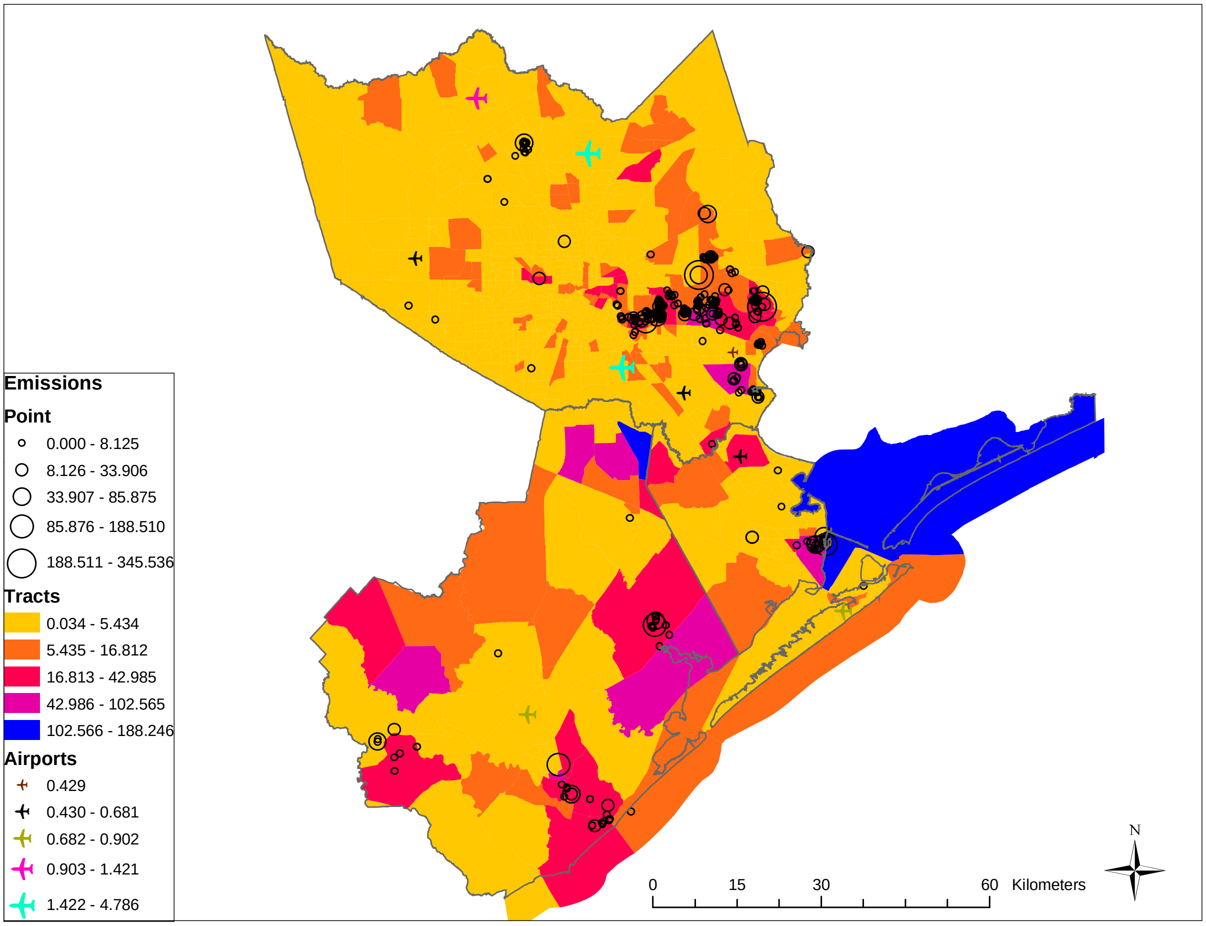


Figure 23. 1990 point (circles), airports (airplane symbols), and spatially allocated total tract (nonpoint plus nonroad) emissions. Emissions are in tons.

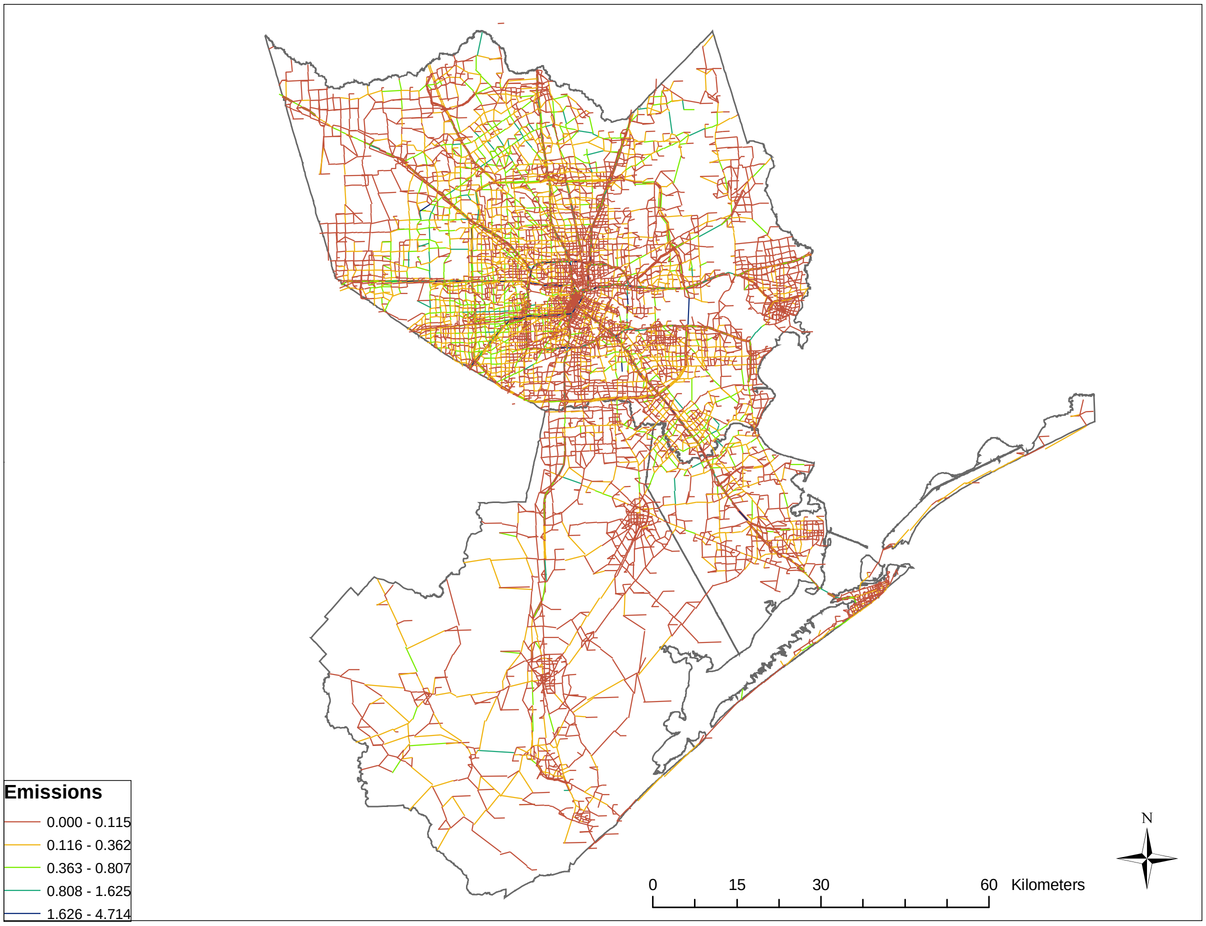


Figure 24. 1990 onroad emissions in tons.

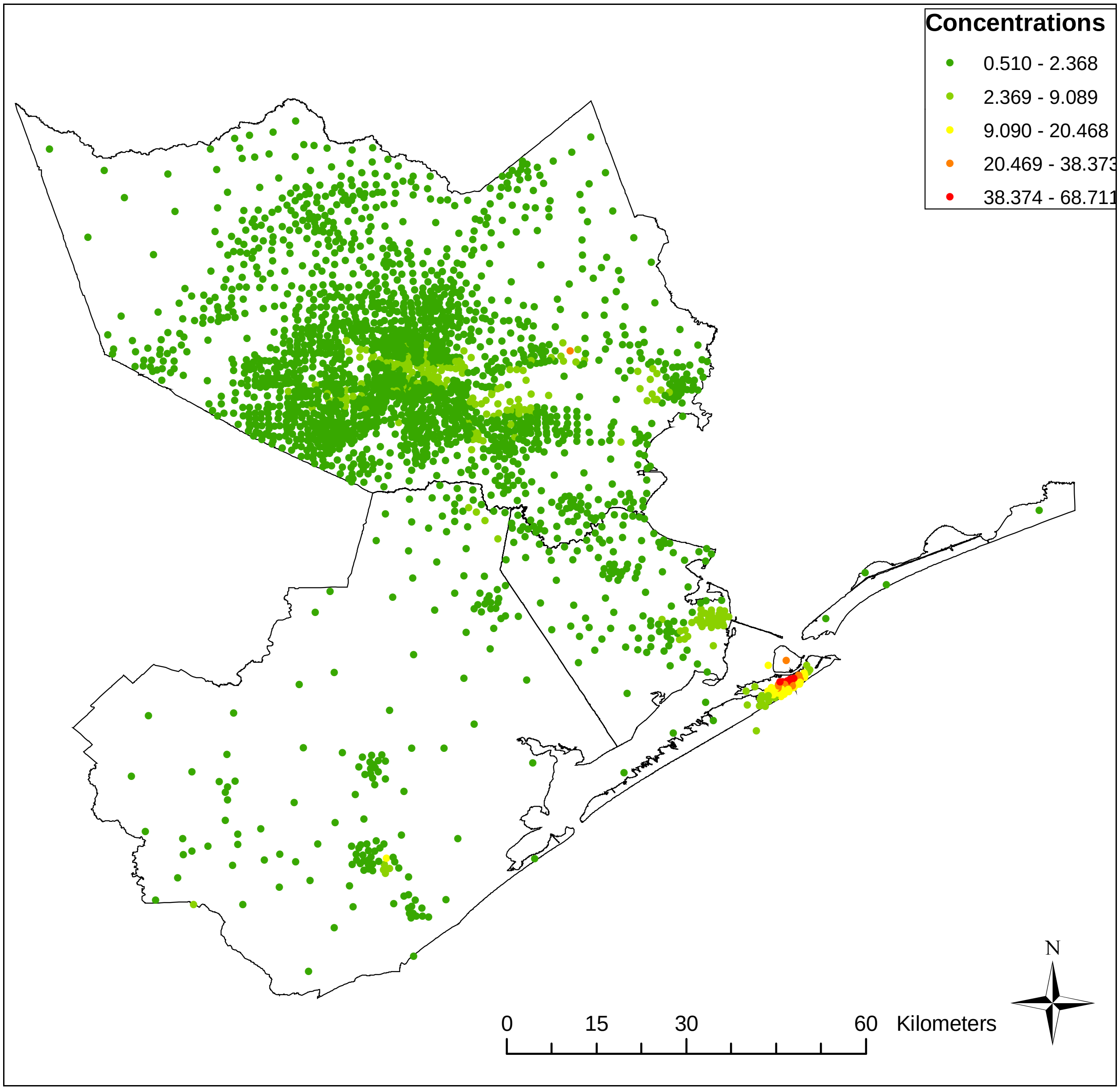


Figure 25. Annual average AERMOD concentrations (μg per m^3) at the census tract block level (with background included) for year 1990.

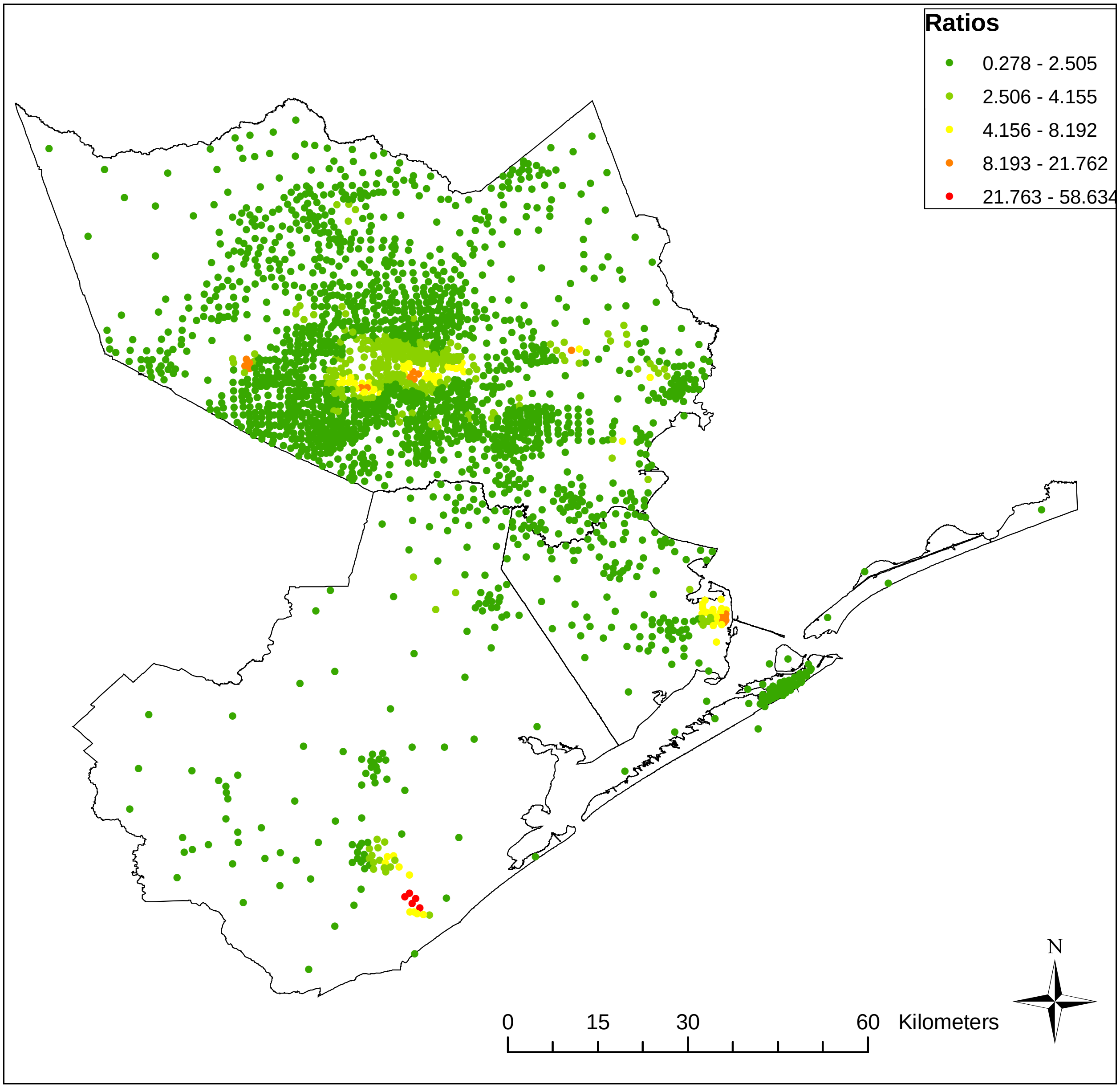


Figure 26. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2000. Concentrations are total concentrations with background included.

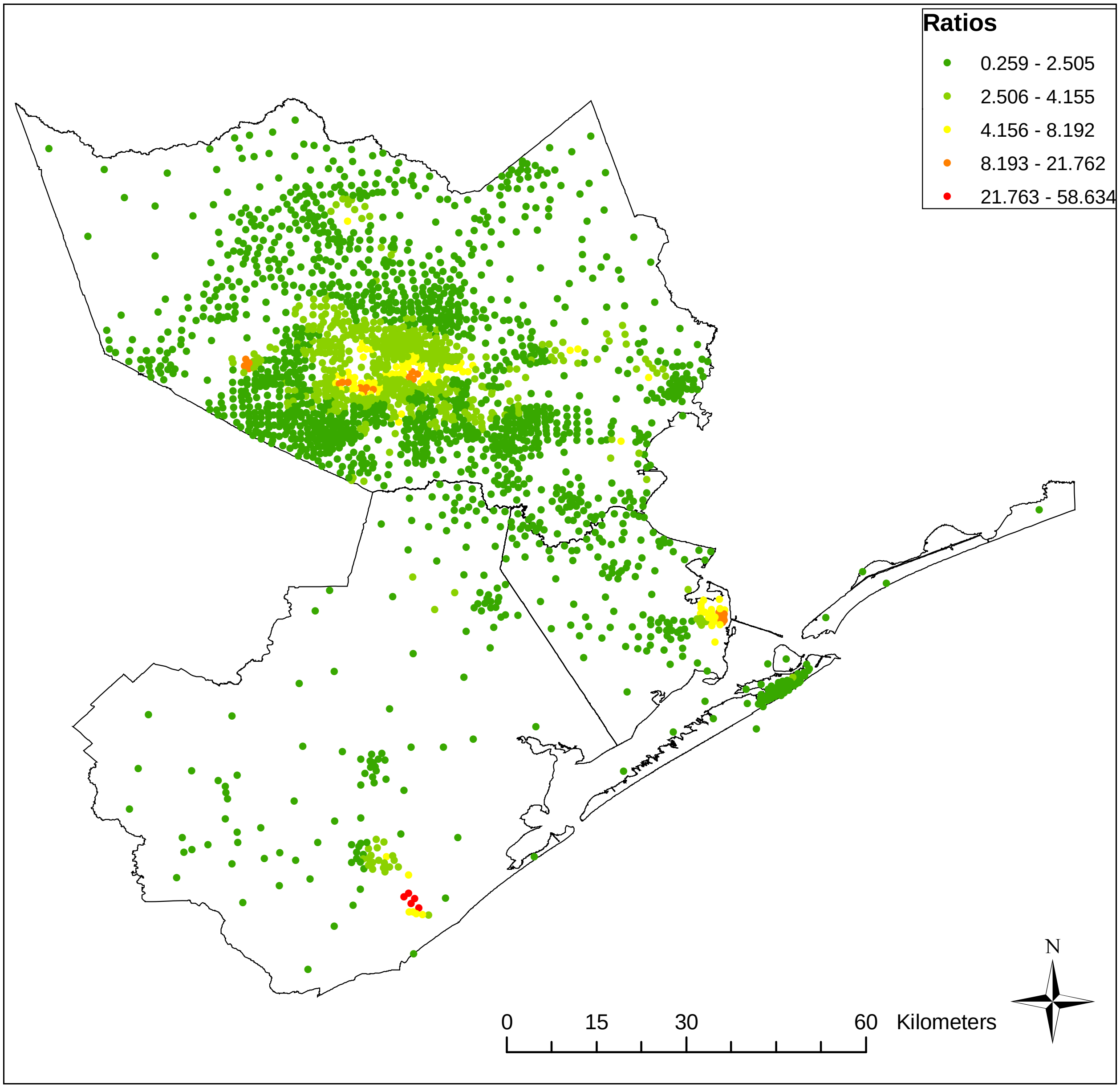


Figure 27. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2010. Concentrations are total concentrations with background included.

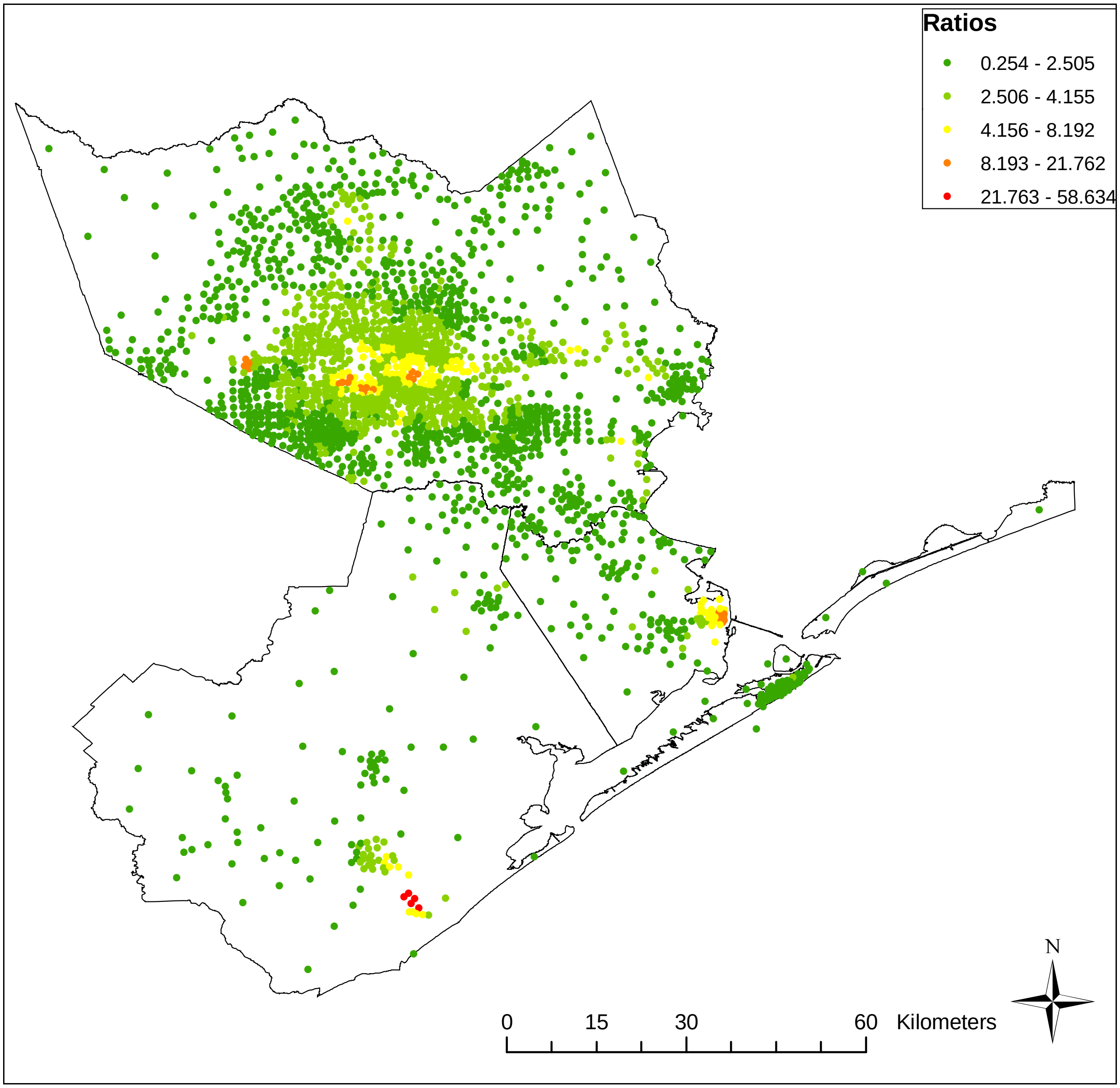


Figure 28. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2020. Concentrations are total concentrations with background included.

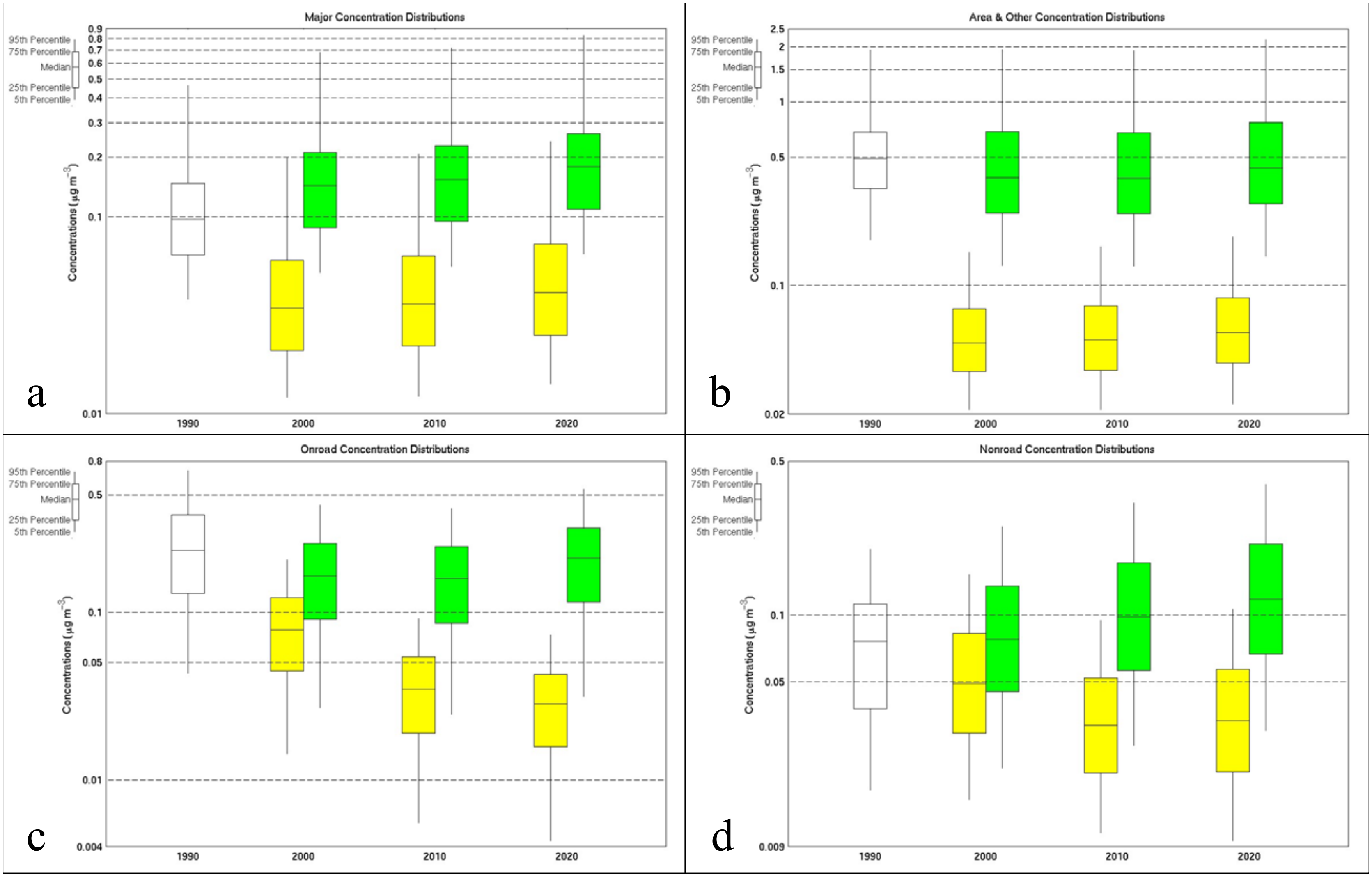


Figure 29. Box and whisker plots of block group level concentration (µg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. The white box denotes 1990, yellow boxes denote the CAA results and green boxes denote the NON-CAA results.

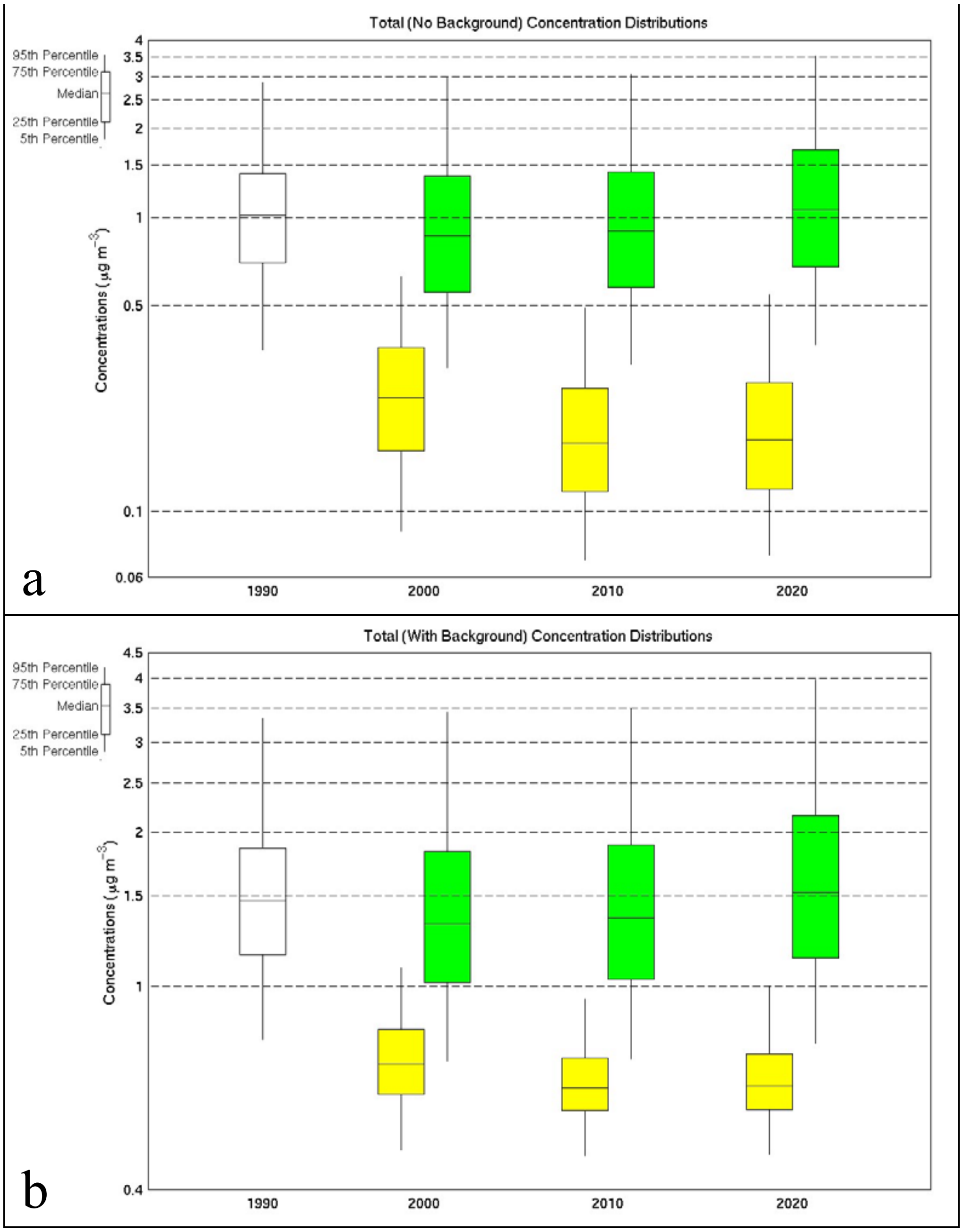


Figure 30. Box and whisker plots of block group level total concentration (μg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) without background and b) with background. Box colors are the same as in Figure 29.

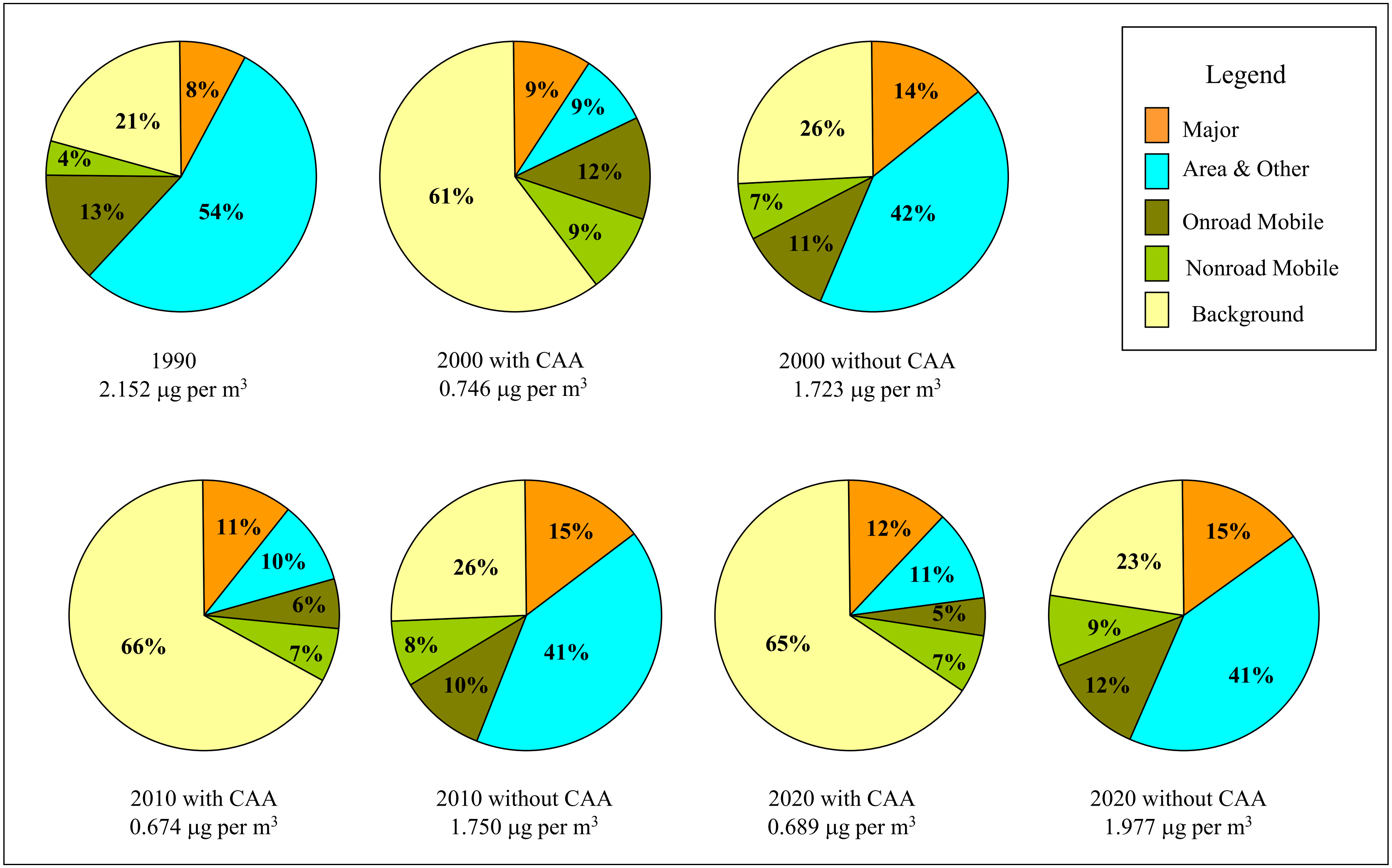


Figure 31. Percent contributions of each source category to the annual average AERMOD total concentrations averaged over all receptors. Total concentrations are given below each chart.

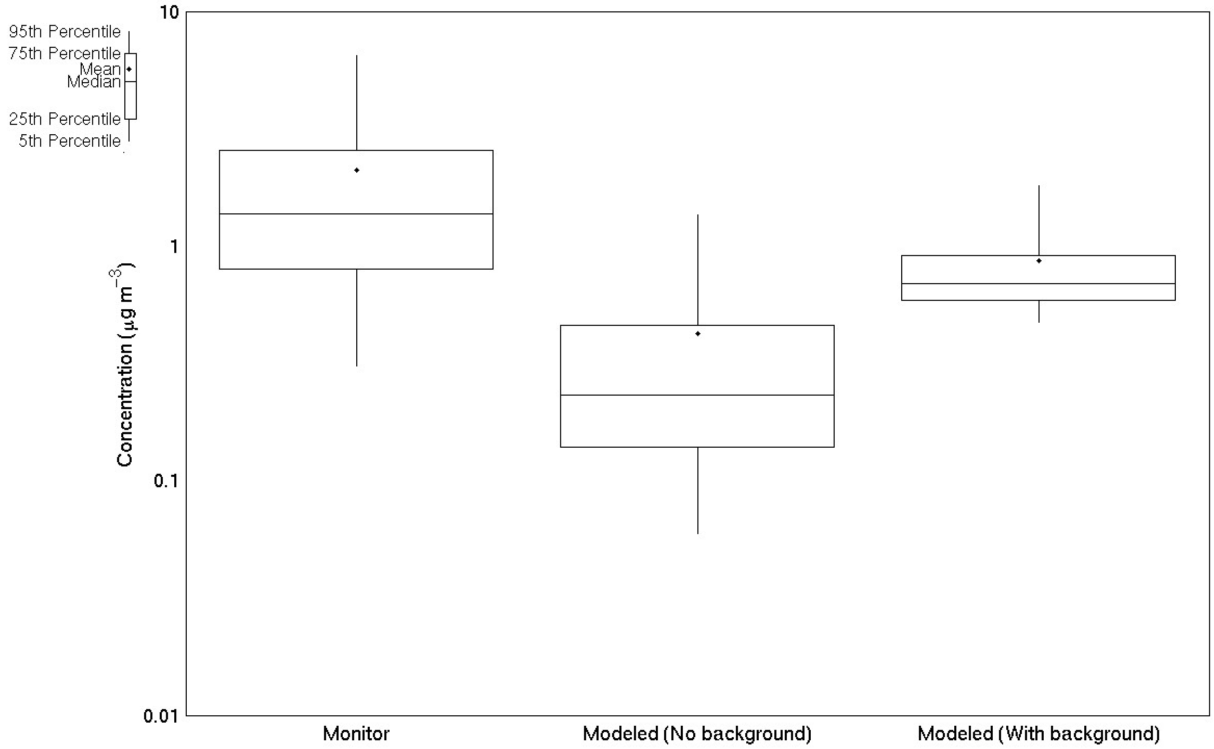


Figure 32. Box and whisker plots of daily monitor, modeled (without background), and modeled (with background) concentrations (µg per m^3). For the AERMOD model, only days with monitor values are included in the calculations.

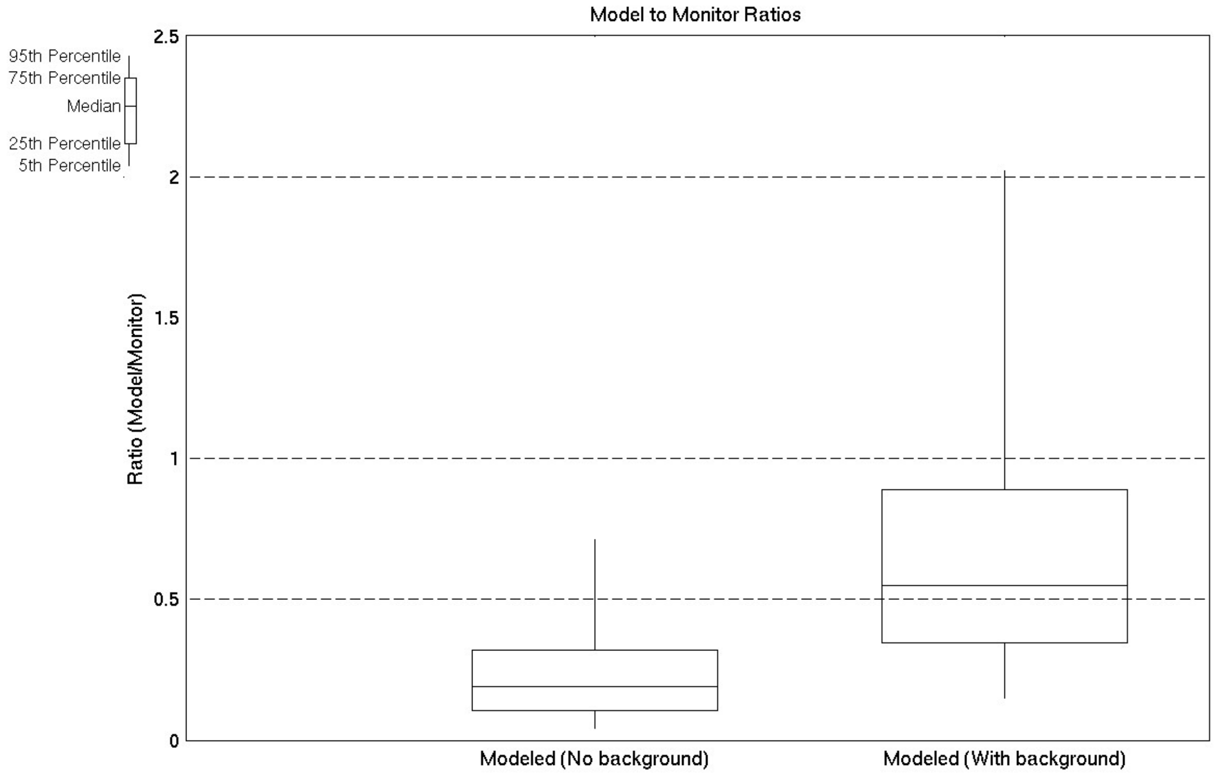


Figure 33. Ratios of daily 2000 CAA AERMOD modeled (with and without background) concentrations to observed monitor concentrations. For the AERMOD model, only days of monitor observations are included in the calculations.

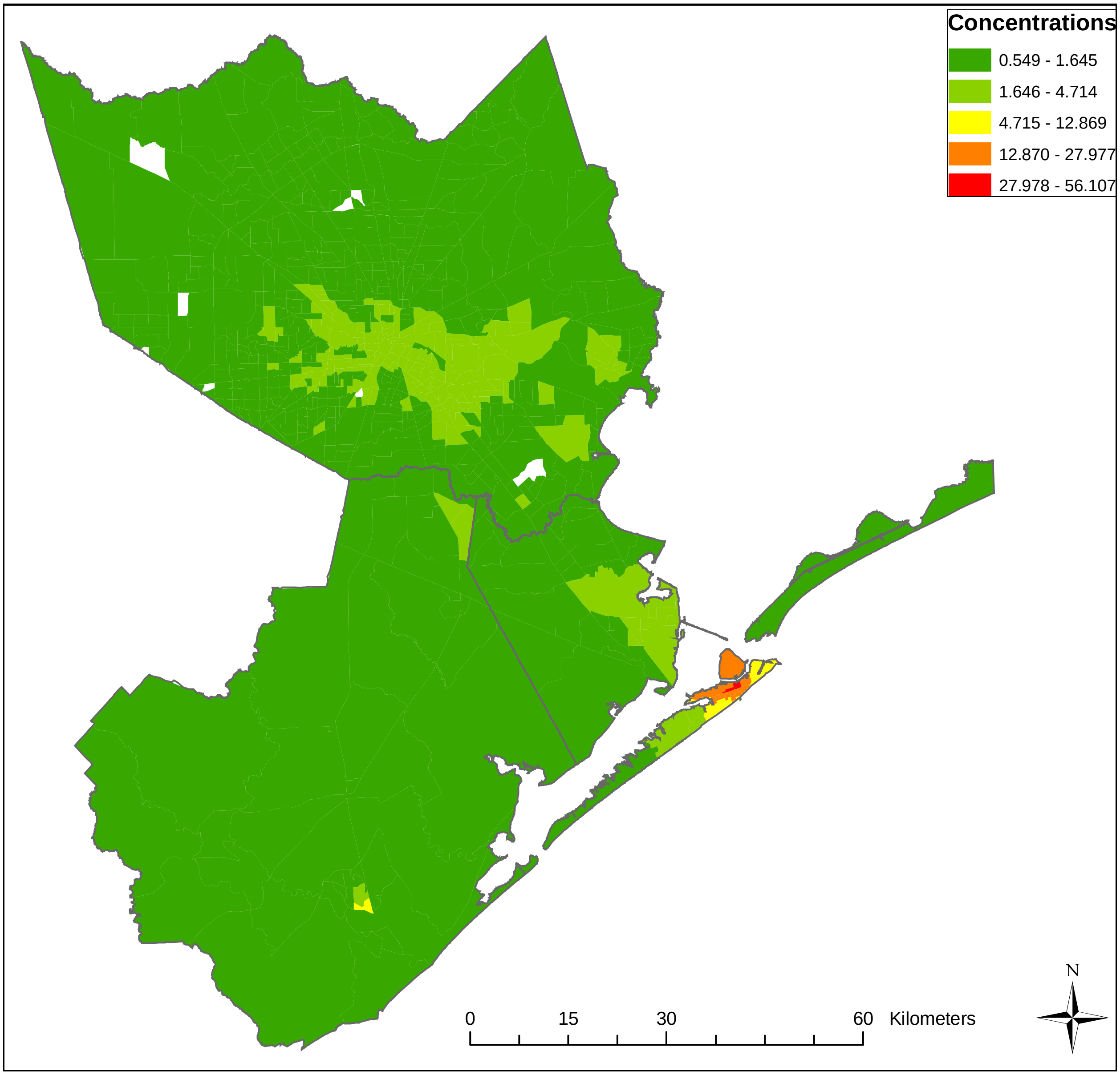


Figure 34. Annual average HAPEM6 concentrations (µg per m^3) at the census tract level (with background included) for year 1990.

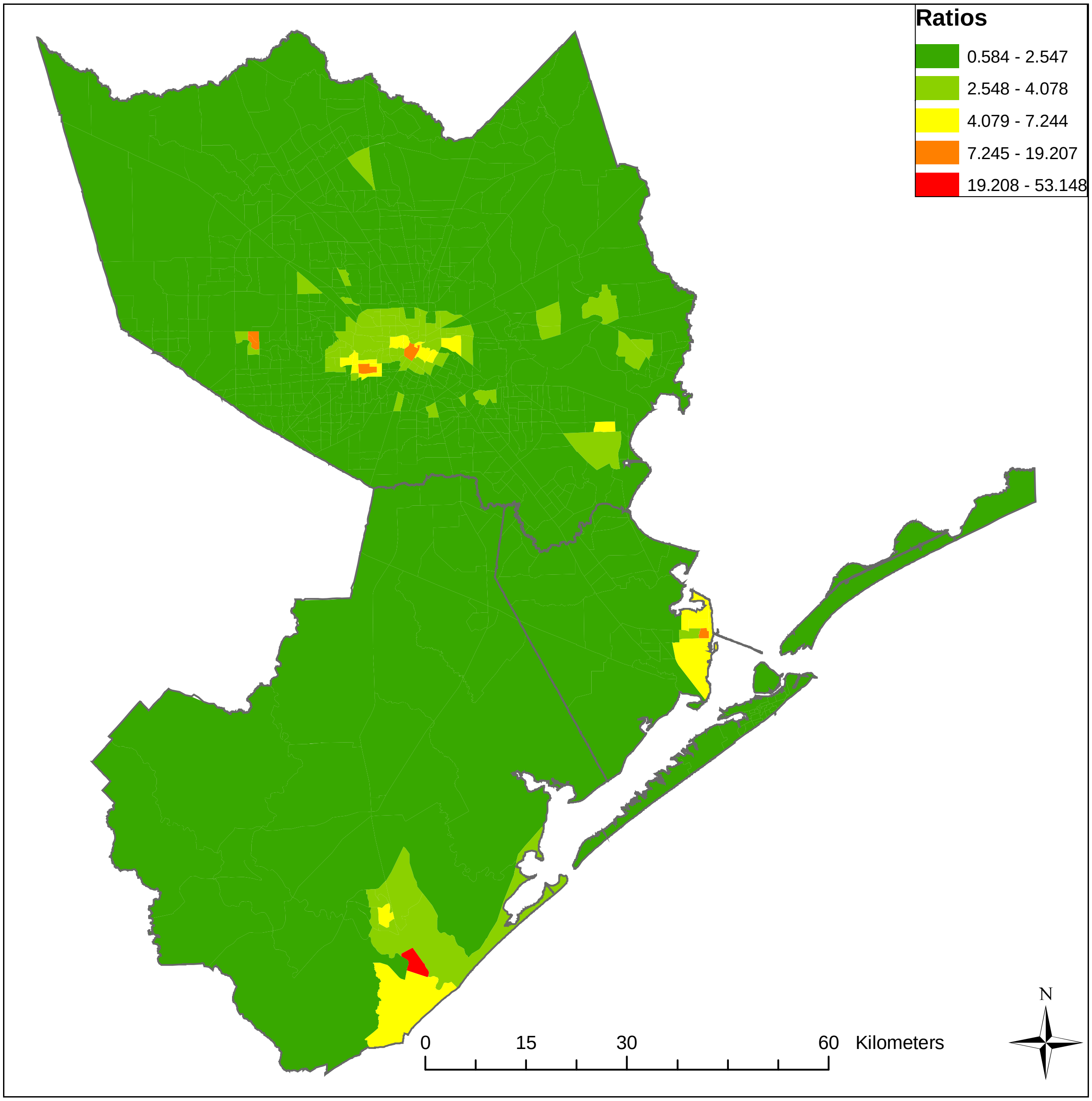


Figure 35. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2000. Concentrations are total concentrations with background included.

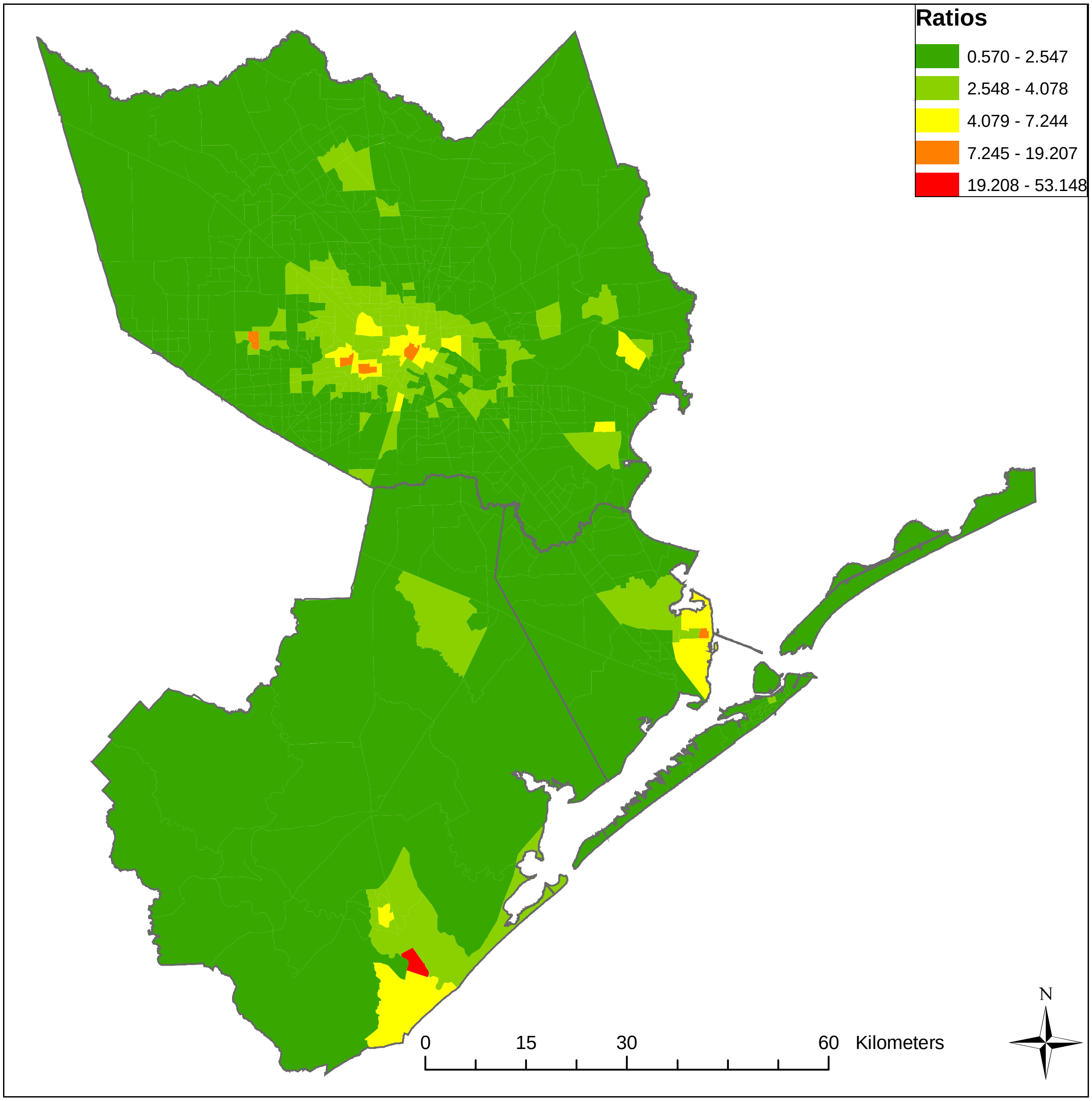


Figure 36. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2010. Concentrations are total concentrations with background included.

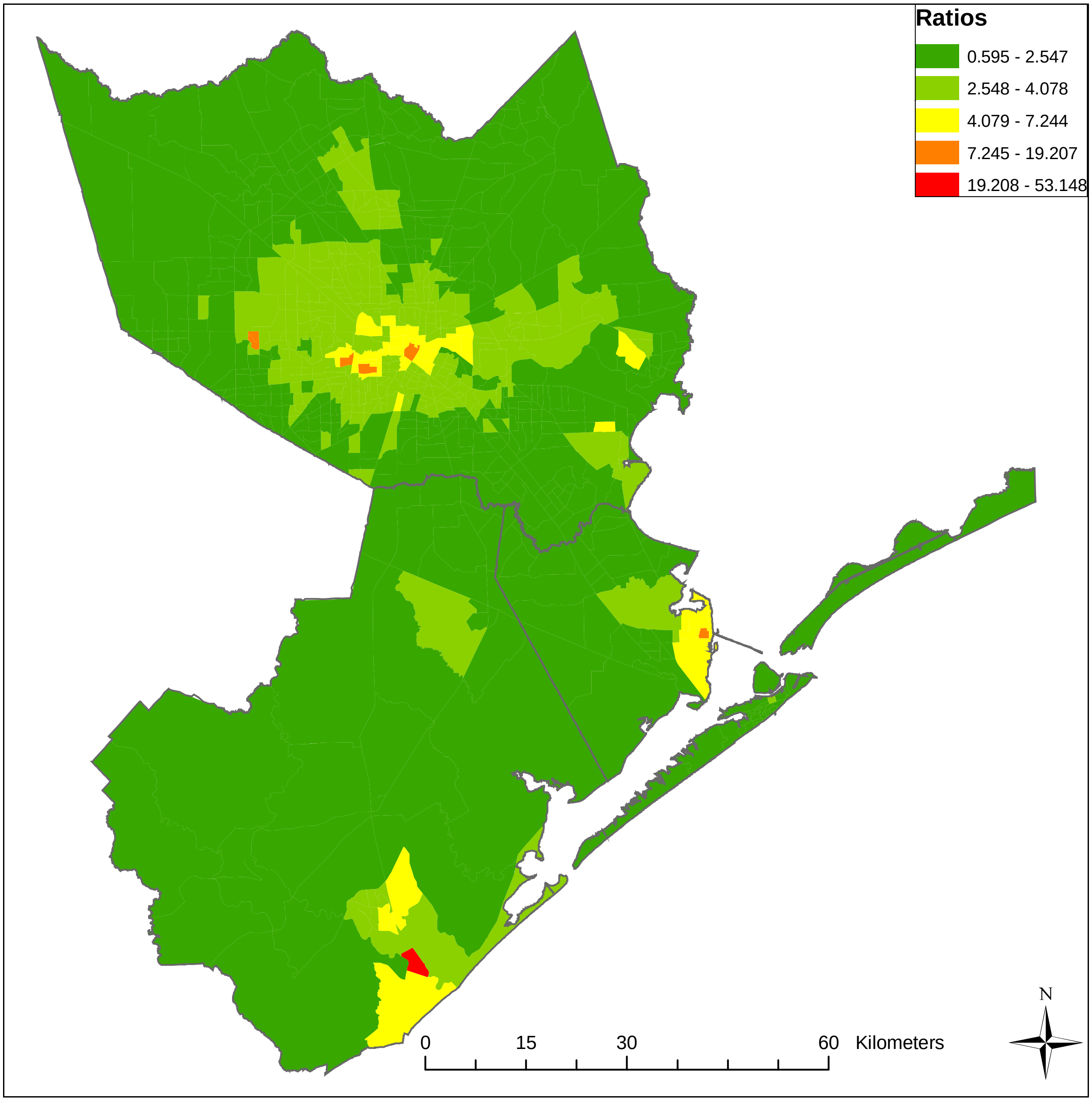


Figure 37. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2020. Concentrations are total concentrations with background included.

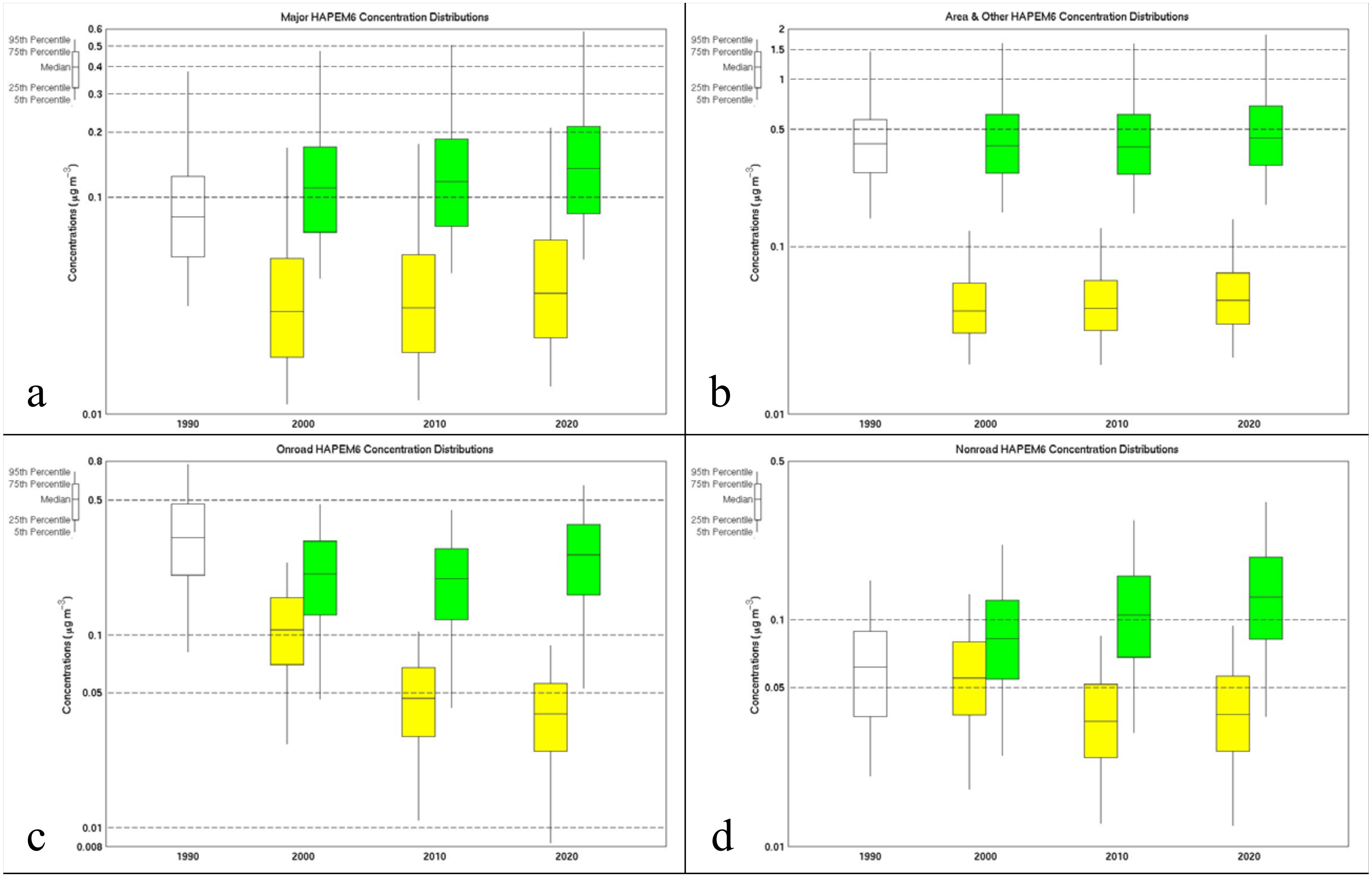


Figure 38. Box and whisker plots of tract level HAPEM6 concentration (µg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. Box colors are the same as in Figure 29.

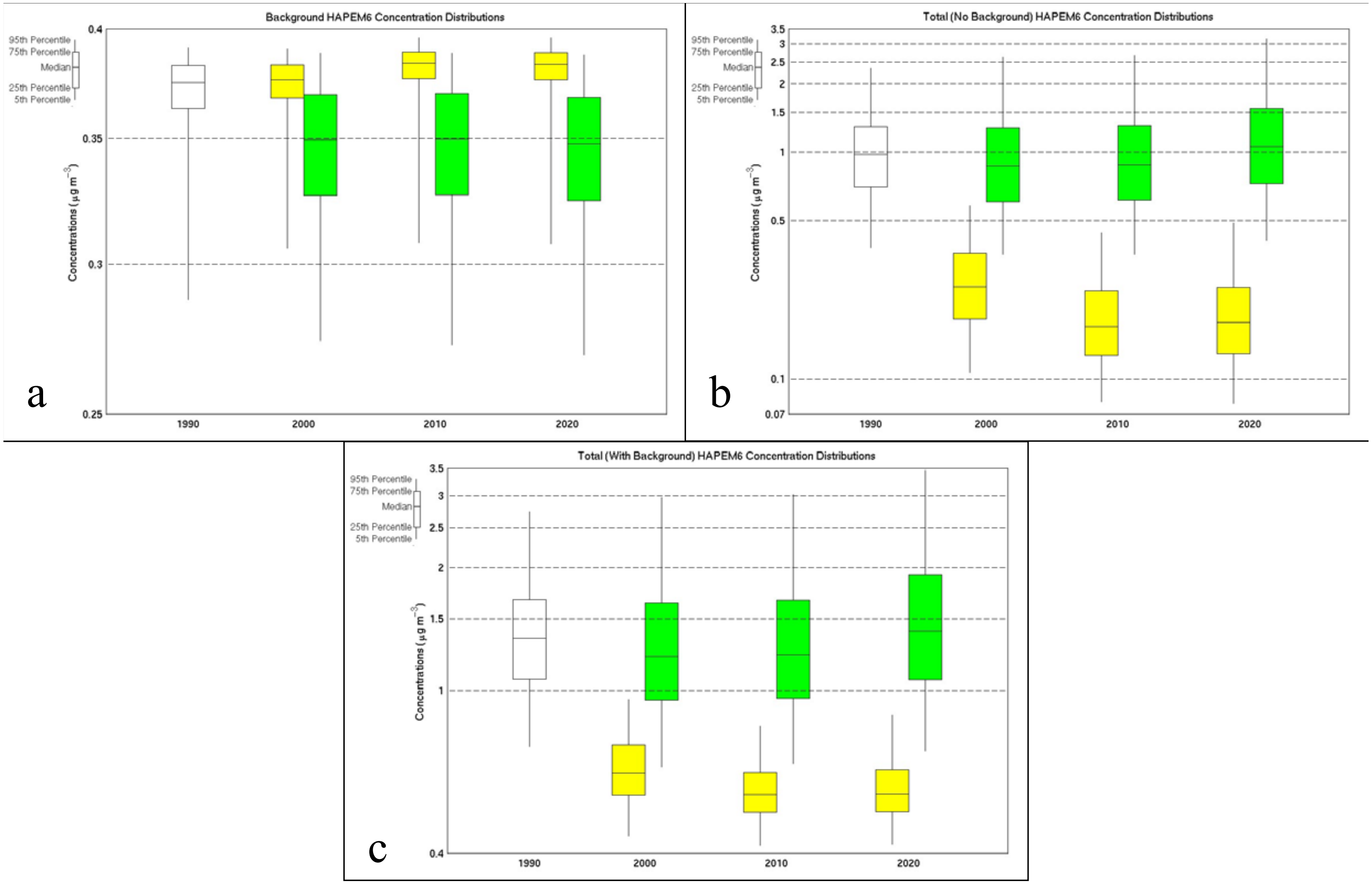


Figure 39. Box and whisker plots of tract level HAPEM6 concentration (µg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) background, b) all sources not including background, and c) all sources including background included. Box colors are the same as in Figure 29.

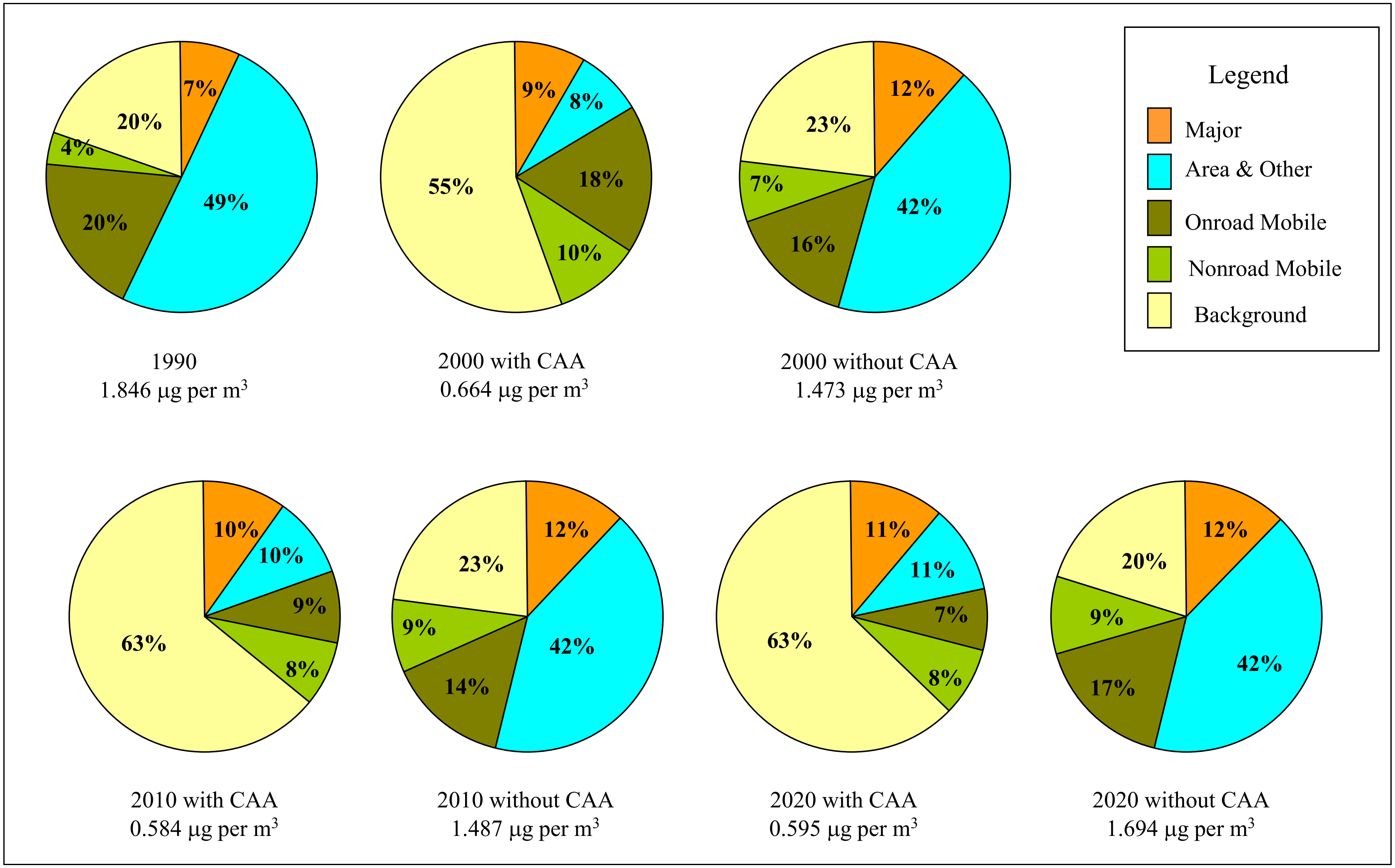


Figure 40. Percent contributions of each source category to the annual average HAPEM6 total concentrations averaged over all receptors. Total concentrations are given below each chart.

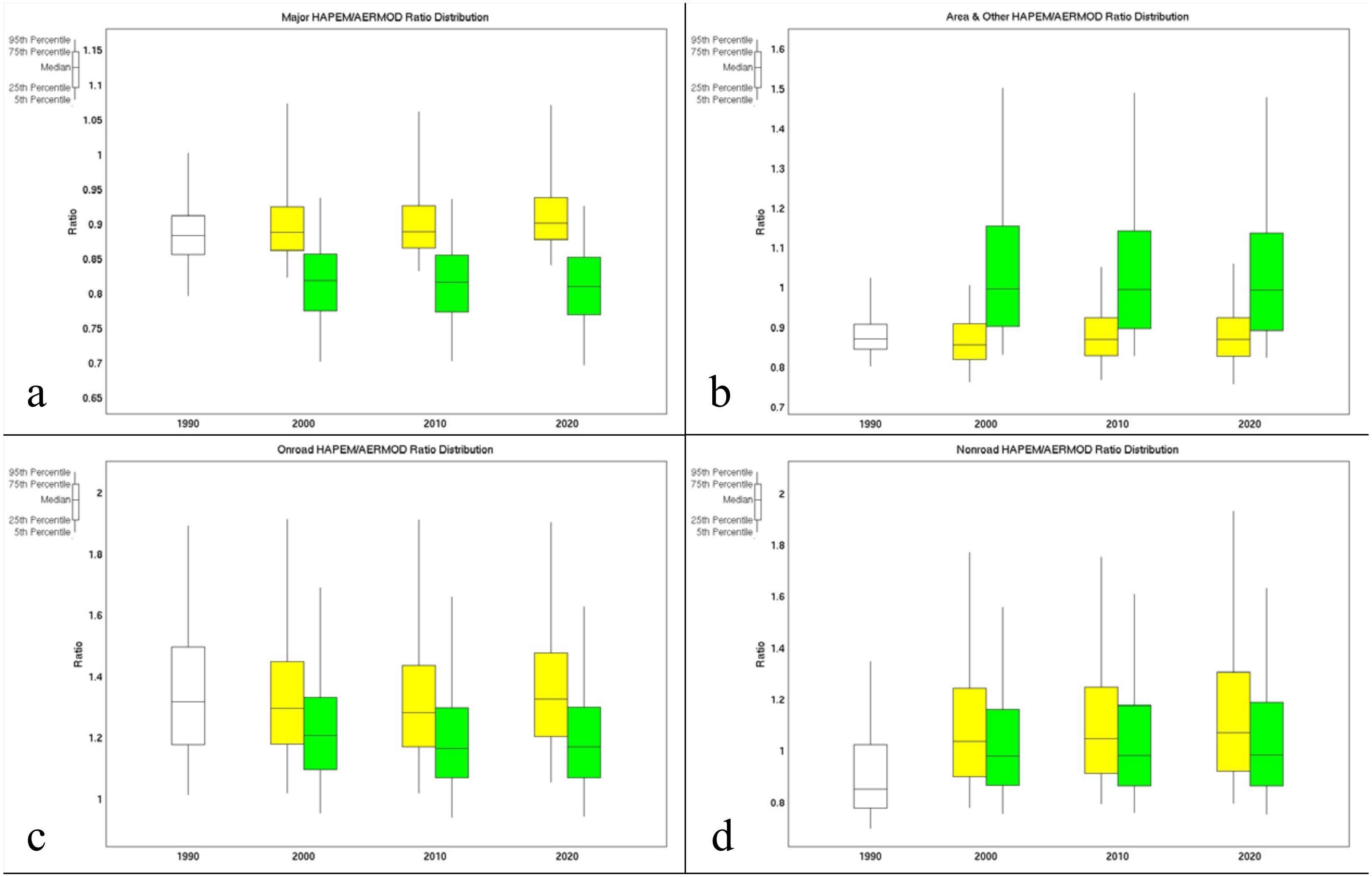


Figure 41. Box and whisker plots of tract level HAPEM6/AERMOD ratio distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. Box colors are the same as in Figure 29.

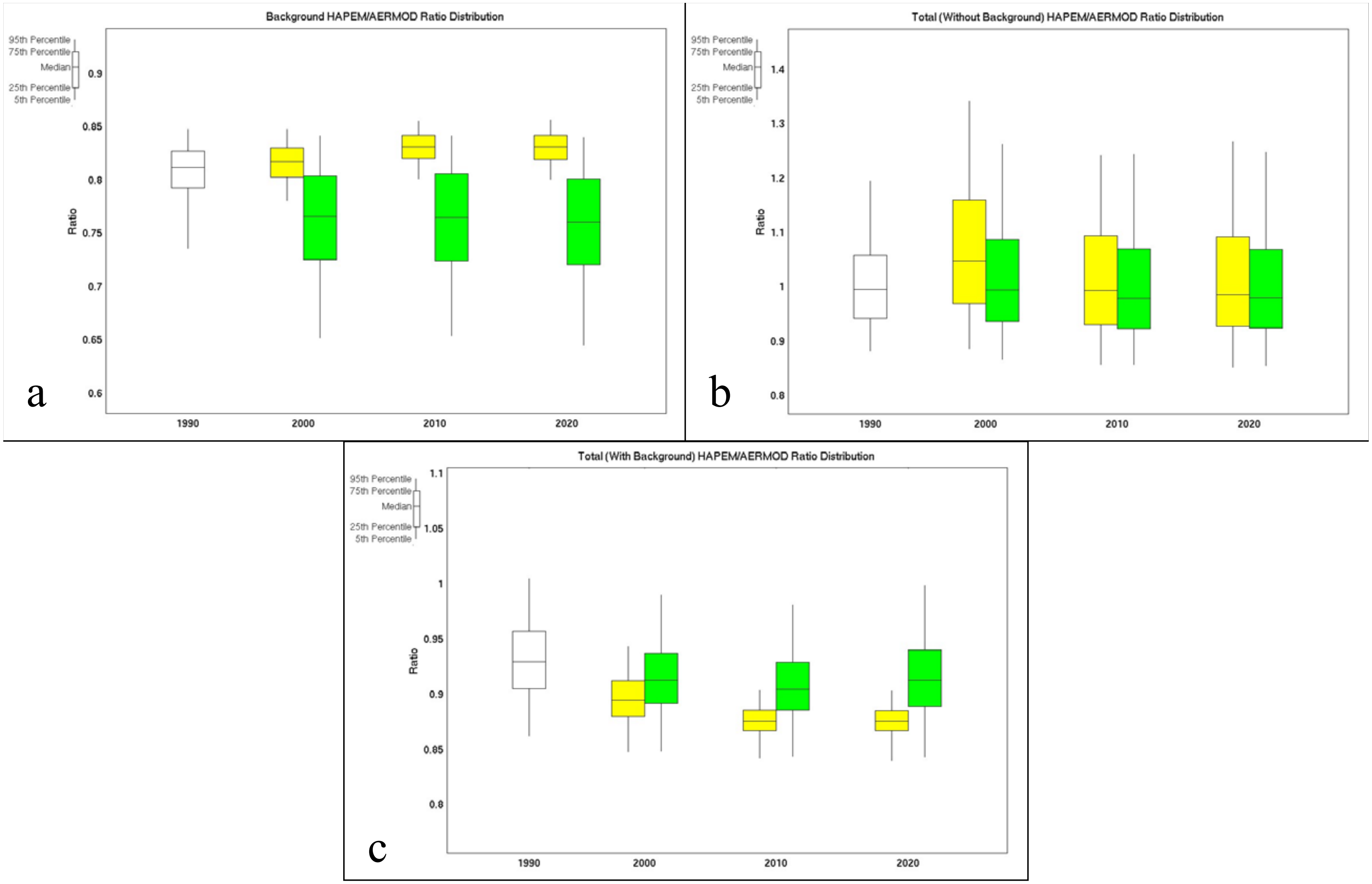


Figure 42. Box and whisker plots of tract level HAPEM6/AERMOD ratio distributions for 1990, 2000, 2010, and 2020 for a) background b) all sources not including background, and c) all sources. Box colors are the same as in Figure 29.

Attachment B-1: Modification of EMS-HAP ancillary files

Before processing in EMS-HAP, ancillary or support files for EMS-HAP needed to be updated for the Houston domain. This included updating the surrogate cross reference file need for spatial allocation of county level emissions (See Chapter 9 of the EMS-HAP User's Guide); updating the temporal allocation factor file to temporally allocate annual emissions to season, day of week, and hour of day; creating tract polygons for the 1990 tracts; and creating new airport allocation factor files for each year as well as generating the source parameters of the airports. Detailed here is the development of the temporal and surrogate cross-reference files.

The temporal allocation file to be used (based off the current version used in EMS-HAP) was checked to see if any SCC codes from the point, nonpoint, and nonroad inventories would be added. SCC codes already in the file would use their current factors. The same file would be used for all inventories. Table A-1 lists the SCC codes that were added to the temporal allocation file for major, area & other, and nonroad emissions. Tables A4-A7 list the assignment of onroad SCC codes to the four new SCC codes. Most SCC codes were assigned factors of related SCC codes. Some were assigned factors based on their temporal allocation factors used for the 1999 NATA (U.S. EPA, 2004).

The surrogate cross reference file lists SCC codes for the nonpoint, nonroad mobile, and onroad mobile (if not using link based emissions) and spatial surrogates used for spatial allocation of county level emissions to census tracts. For example, industrial related emissions may be allocated to tracts using industrial land area in each tract. The nonpoint and nonroad inventories for each year were checked to see if any SCC codes in the inventories were not in the current surrogate cross reference files as used for the 1999 NATA [See Table C-6 of EMS-HAP User-s Guide (U.S. EPA, 2004) for list of SCC codes in current file]. SCC codes already in the files would use their current surrogate assignments. There were two surrogate files to consider. The first was based on the 1990 census (U.S. EPA, EMS-HAP UG V2) and the other based on the 2000 census (U.S. EPA, 2004). The 1990 base nonpoint and nonroad inventories were compared against the 1990 tract surrogates while the other inventories for 2000, 2010, and 2020 were compared against the 2000 census surrogates. The 1990 census data would be used for the 1990 inventories while all other years would use the 2000 census data. It was found that several SCC codes would need surrogate assignments. Those SCC codes could be assigned a surrogate based on a related SCC code already in the surrogate file or files. Table A-2 lists the SCC codes that needed assignments in the 1990 inventories and Table A-3 list those needed for the other years.

Table B-1. SCC codes to add to the temporal allocation factor file.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
30107101	Industrial Processes; Chemical Manufacturing; Hydrogen; Reformers	Point	30107001	Industrial Processes; Chemical Manufacturing; Inorganic Chemical Manufacturing (General); Fugitive Leaks	Related SCC code
30112558	Industrial Processes; Chemical Manufacturing; Chlorine Derivatives; Chloromethanes via MH & MCC Processes: Chloroform Condenser	Point	30112555	Industrial Processes; Chemical Manufacturing; Chlorine Derivatives; Vinylidene Chloride: Fugitive Emissions	Related SCC codes use same profiles
30130115	Industrial Processes; Chemical Manufacturing; Chlorobenzene; Atmospheric Distillation Vents	Point	30130110	Industrial Processes; Chemical Manufacturing; Chlorobenzene; Catalyst Incineration	Related SCC codes use same profiles
30301582	Industrial Processes; Primary Metal Production; Integrated Iron and Steel Manufacturing (See also 3-03-008 & 3-03-009); Miscellaneous Combustion Sources: Boilers	Point	30301499	Industrial Processes; Primary Metal Production; Barium Ore Processing; Other Not Classified	Similar SCC codes use same profiles
30600518	Industrial Processes; Petroleum Industry; Wastewater Treatment; Petroleum Refinery Wastewater System: Weir	Point	30600517	Industrial Processes; Petroleum Industry; Wastewater Treatment; Petroleum Refinery Wastewater System: Non-aerated Impoundment	Related SCC codes use same profiles
30601601, 30601604	Industrial Processes; Petroleum Industry; Catalytic Reforming Unit	Point	2306000000	Industrial Processes; Petroleum Refining: SIC 29; All Processes; Total	2306000000 is most general of refining SCC codes. Using the profile associated with 2306000000 will yield a flat profile, all hours the same.

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
30601701	Industrial Processes; Petroleum Industry; Catalytic Hydrotreating Unit; General	Point	2306000000	Industrial Processes; Petroleum Refining: SIC 29; All Processes; Total	2306000000 is most general of refining SCC codes. Using the profile associated with 2306000000 will yield a flat profile, all hours the same.
30602001	Industrial Processes; Petroleum Industry; Crude Unit Atmospheric Distillation; General	Point	2306000000	Industrial Processes; Petroleum Refining: SIC 29; All Processes; Total	2306000000 is most general of refining SCC codes. Using the profile associated with 2306000000 will yield a flat profile, all hours the same.
30602101	Industrial Processes; Petroleum Industry; Light Ends Fractionation Unit; General	Point	2306000000	Industrial Processes; Petroleum Refining: SIC 29; All Processes; Total	2306000000 is most general of refining SCC codes. Using the profile associated with 2306000000 will yield a flat profile, all hours the same.

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
30630007	Industrial Processes; Petroleum Industry; Re-refining of Lube Oils and Greases; Finished Product Storage Tank	Point	30630005	Industrial Processes; Petroleum Industry; Re-refining of Lube Oils and Greases; Waste Oil Still Vent	Related codes 30630005 and 3630006 use same profile
31000220	Industrial Processes; Oil and Gas Production; Natural Gas Production; All Equipt Leak Fugitives (Valves, Flanges, Connections, Seals, Drains	Point	31000223	Industrial Processes; Oil and Gas Production; Natural Gas Production; Relief Valves	Use profile of related SCC code
40400140	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Bulk Terminals; Specify Liquid: Standing Loss - Ext. Float Roof Tank w/ Second'y Seal	Point	40400141	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Bulk Terminals; Gasoline RVP 13: Standing Loss - Ext. Floating Roof w/ Secondary Seal	Use profile of related SCC code
40400252	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Bulk Plants; Miscellaneous Losses/Leaks: Vapor Collection Losses	Point	40400251	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Bulk Plants; Valves, Flanges, and Pumps	Use profile of related SCC code
40400312, 40400321, 40400324, 40400340	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Oil and Gas Field Storage and Working Tanks	Point	40400301	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Oil and Gas Field Storage and Working Tanks; Fixed Roof Tank: Breathing Loss	Use profile of related SCC code
40400498	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Petroleum Products - Underground Tanks; Specify Liquid: Working Loss	Point	40400497	Petroleum and Solvent Evaporation; Petroleum Liquids Storage (non-Refinery); Petroleum Products - Underground Tanks; Specify Liquid: Breathing Loss	Use profile of related SCC code
40600131, 40600136, 40600141, 40600144, 40600170, 40600197	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Tank Cars and Trucks	Point	40600172	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Tank Cars and Trucks; Transit Losses - LPG: Loaded with Fuel	use profile of related SCC code

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
40600231, 40600234, 40600240, 40600244, 40600246, 40600249,	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Marine Vessels	Point	40600241	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Marine Vessels; Gasoline: Tanker Ship - Ballasting	use profile of related SCC code
40600259, 40600298, 40600299	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Marine Vessels	Point	40600256	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Marine Vessels; Kerosene: Transit Loss	use profile of related SCC code
40688801, 40688802	Petroleum and Solvent Evaporation; Transportation and Marketing of Petroleum Products; Fugitive Emissions	Point			Use a uniform or flat profile as done for ASPEN. Also other 406XXXXXX SCC codes use uniform profile.
40700816, 40700898	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Alcohols	Point			Use a uniform or flat profile as done for ASPEN. Also other 407XXXXXX SCC codes use uniform profile.
40701611, 40701613, 40701614	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Alkanes (Paraffins)	Point	40701610	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Alkanes (Paraffins); Pentadecane: Working Loss	use profile of related SCC code
40702097	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Alkenes (Olefins); Specify Olefin: Breathing Loss	Point	40702004	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Alkenes (Olefins); Heptenes - General: Working Loss	use profile of related SCC code

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
40703602, 40703610, 40703614, 40703697, 40703698	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Aromatics	Point	40703611	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Aromatics; Methyl Styrene: Breathing Loss	use profile of related SCC code
40704898	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Ethers; Specify Ether: Working Loss	Point	40704424	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Esters; i-Butyl-i-Butyrate: Working Loss	40704XXX SCC codes have same profile
40706097	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Halogenated Organics; Specify Halogenated Organic: Breathing Loss	Point	40706016	Petroleum and Solvent Evaporation; Organic Chemical Storage; Fixed Roof Tanks - Halogenated Organics; Ethylene Dibromide: Working Loss	40706XXX SCC codes have same profile
40717697, 40717698	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Alkanes (Paraffins)	Point	40717606	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Alkanes (Paraffins); n-Pentane: Withdrawal Loss	407176XXX codes have same profile
40718097, 40718098	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Alkenes (Olefins)	Point	40718010	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Alkenes (Olefins); Cyclopentene: Withdrawal Loss	407180XX codes have same profile
40719601, 40719602	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Aromatics; Benzene: Standing Loss	Point	40719615	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Aromatics; Toluene: Standing Loss	407196XX codes have same profile
40720897	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Ethers; Specify Ether: Standing Loss	Point	40720802	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Ethers; Ethyl Ether: Withdrawal Loss	4072XXXX codes have same profile
40721205	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Glycol Ethers; Carbitol: Standing Loss	Point	40720802	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Ethers; Ethyl Ether: Withdrawal Loss	4072XXXX codes have same profile

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
40729698	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Miscellaneous; Specify In Comments: Working Loss	Point	40729697	Petroleum and Solvent Evaporation; Organic Chemical Storage; Floating Roof Tanks - Miscellaneous; Specify In Comments: Breathing Loss	4072XXXX codes have same profile
40799997, 40799998	Petroleum and Solvent Evaporation; Organic Chemical Storage; Miscellaneous	Point	40799999	Petroleum and Solvent Evaporation; Organic Chemical Storage; Miscellaneous; Other Not Classified	Use related SCC code already in file
40899995, 40899997, 40899999	Petroleum and Solvent Evaporation; Organic Chemical Transportation; Specific Liquid	Point			flat profile as done for ASPEN. Also other 408XXXXXXX SCC codes use uniform profile.
49000201, 49000206, 49000299	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Waste Solvent Recovery Operations	Point	49000207	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Waste Solvent Recovery Operations; Distillation Vent	Use related SCC code already in file
49000399	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Rail Car Cleaning; Other Not Classified	Point	49000304	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Rail Car Cleaning; Creosote	Use related SCC code already in file
49090013	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Fuel Fired Equipment; Natural Gas: Incinerators	Point	49090012	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Fuel Fired Equipment; Residual Oil: Incinerators	Use related SCC code already in file
49099998	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Miscellaneous Volatile Organic Compound Evaporation; Identify the Process and Solvent in Comments	Point			flat profile as done for ASPEN
50100403	Waste Disposal; Solid Waste Disposal - Government; Landfill Dump; Area Method	Point	50100402	Waste Disposal; Solid Waste Disposal - Government; Landfill Dump; Fugitive Emissions	Use related SCC code already in file

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
50100701, 50100702, 50100703, 50100704, 50100707, 50100710, 50100720, 50100732, 50100740, 50100791	Waste Disposal; Solid Waste Disposal - Government; Sewage Treatment	Point			flat profile as done for ASPEN
50300801	Waste Disposal; Solid Waste Disposal - Industrial; Treatment, Storage, Disposal/TSDF; Surface Impoundment: Fugitive Emissions	Point	50300899	Waste Disposal; Solid Waste Disposal - Industrial; Treatment, Storage, Disposal/TSDF; General: Fugitive Emissions	Use related SCC code already in file
64420033	MACT Source Categories; Cellulose-based Resins; Carboxymethylcellulose Production; Product Finishing: Purification/Extraction				flat profile as done for ASPEN
2302002100, 2302002200	Industrial Processes; Food and Kindred Products: SIC 20; Commercial Charbroiling	Nonpoint	2302002000	Industrial Processes; Food and Kindred Products: SIC 20; Commercial Charbroiling; Total	Use related SCC code already in file
2310001000	Industrial Processes; Oil and Gas Production: SIC 13;	Nonpoint	2310000000	Industrial Processes; Oil and Gas Production: SIC 13; All Processes; Total: All Processes	Use related SCC code already in file
2501055120	Storage and Transport; Petroleum and Petroleum Product Storage; Bulk Stations/Terminals: Breathing Loss	Nonpoint	2501050000	Storage and Transport; Petroleum and Petroleum Product Storage; Bulk Stations/Terminals: Breathing Loss; Total: All Products	Use related SCC code already in file

Table B-1. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
49000199	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Solvent Extraction Process; Other Not Classified	Nonpoint	49000102	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Solvent Extraction Process; Methyl Ethyl Ketone	The 49000XXX SCC codes already in file use same profile.
50300601	Waste Disposal; Solid Waste Disposal - Industrial; Landfill Dump; Waste Gas Flares	Nonpoint	50300203	Waste Disposal; Solid Waste Disposal - Industrial; Open Burning; Auto Body Components	All 50300XXX SCC codes already in file have same profiles.
2267005055	Mobile Sources; LPG; Agricultural Equipment; Other Agricultural Equipment	Nonroad	2267005050	Mobile Sources; LPG; Agricultural Equipment; Hydro-power Units	2267005XXX SCC codes in file have same profiles.
2268002081	Mobile Sources; CNG; Construction and Mining Equipment; Other Construction Equipment	Nonroad	2268002000	Mobile Sources; CNG; Construction and Mining Equipment; All	2268002XXX SCC codes in file have same profile
2268003020, 2268003040, 2268003070	Mobile Sources; CNG; Industrial Equipment;	Nonroad	2268003030	Mobile Sources; CNG; Industrial Equipment; Sweepers/Scrubbers	2268003XXX SCC codes in file have same profiles
2268005050, 2268005055	Mobile Sources; CNG; Agricultural Equipment	Nonroad	2268005060	Mobile Sources; CNG; Agricultural Equipment; Irrigation Sets	2268005XXX codes in file have similar profiles
2268006015	Mobile Sources; CNG; Commercial Equipment; Air Compressors	Nonroad	2268006010	Mobile Sources; CNG; Commercial Equipment; Pumps	2268006XXX codes in file have same profile

Table B-2. Surrogates assigned to SCC codes in 1990 inventories.

SCC code(s)	Description	Inventory	Surrogate	Reasons for surrogate choice
10200401, 10200501, 10200601	External Combustion Boilers; Industrial	Nonpoint	Industrial land	Other industrial boiler SCC codes use industrial land
10300701	External Combustion Boilers; Commercial/Institutional; Process Gas; POTW Digester Gas-fired Boiler	Nonpoint	Inverse pop. density using tract land area	Other Commercial/Institutional POTW gas fired boilers used Inverse pop. density
20100201	Internal Combustion Engines; Electric Generation; Natural Gas; Turbine	Nonpoint	Industrial land	20100202, a related SCC uses industrial land
2104008002, 2104008010	Stationary Source Fuel Combustion; Residential; Wood	Nonpoint	Population	Residential wood burning SCC codes in surrogate file use population.
2301040000	Industrial Processes; Chemical Manufacturing: SIC 28; Fugitive Emissions from Synthetic Organic Chem Manuf (NAPAP cat. 102); Total	Nonpoint	Industrial land	Industrial processes use industrial land
2460400000	Solvent Utilization; Miscellaneous Non-industrial: Consumer and Commercial; All Automotive Aftermarket Products; Total: All Solvent Types	Nonpoint	Population	In 1999 NATA, 2460400000 used population.
2630020000	Waste Disposal, Treatment, and Recovery; Wastewater Treatment; Public Owned; Total Processed	Nonpoint	Inverse pop. density using tract land area	2630000000 in 1990 surrogates uses inverse population density.
31000299	Industrial Processes; Oil and Gas Production; Natural Gas Production; Other Not Classified	Nonpoint	Industrial land	Oil and natural gas production SCC codes in surrogate file used industrial land
49000199	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Solvent Extraction Process; Other Not Classified	Nonpoint	Industrial land	Industrial process
50300601	Waste Disposal; Solid Waste Disposal - Industrial; Landfill Dump; Waste Gas Flares	Nonpoint	Inverse pop. density using tract land area	Landfills use inverse population density
2260001010, 2260001030, 2260001060	Mobile Sources; Off-highway Vehicle Gasoline, 2-Stroke; Recreational Equipment	Nonroad	Tract area	2-stroke recreation equipment in surrogate file use tract area

Table B-2. Continued.

SCC code(s)	Description	Inventory	Surrogate	Reasons for surrogate choice
2260002054	Mobile Sources; Off-highway Vehicle Gasoline, 2-Stroke; Construction and Mining Equipment; Crushing/Processing Equipment	Nonroad	Inverse pop. density using total tract area	2-stroke Construction and mining equipment in surrogate file use Inverse pop. density
2260005050	Mobile Sources; Off-highway Vehicle Gasoline, 2-Stroke; Agricultural Equipment; Hydro-power Units	Nonroad	farmland + orchard land	Other 2-stroke ag equipment in surrogate file uses farmland + orchard land
2265001010, 2265001030, 2265001050, 2265001060	Mobile Sources; Off-highway Vehicle Gasoline, 4-Stroke; Recreational Equipment	Nonroad	Tract area	4-stroke recreation equipment in surrogate file use tract area
2267002003, 2267002015, 2267002021, 2267002024, 2267002033, 2267002039, 2267002045, 2267002054, 2267002057, 2267002060, 2267002066, 2267002072, 2267002081	Mobile Sources; LPG; Construction and Mining Equipment	Nonroad	Inverse pop. density using total tract area	Construction and mining equipment in surrogate file use inverse pop. density
2267003010, 2267003020, 2267003030, 2267003040, 2267003050, 2267003070	Mobile Sources; LPG; Industrial Equipment	Nonroad	Industrial land	Industrial equipment in surrogate files uses industrial land
2267004066	Mobile Sources; LPG; Lawn and Garden Equipment; Chippers/Stump Grinders (Commercial)	Nonroad	residential land	Lawn and garden equipment in surrogate files uses residential land
2267005050, 2267005055, 2267005060	Mobile Sources; LPG; Agricultural Equipment	Nonroad	farmland + orchard land	Ag equipment in surrogate file uses farmland + orchard land
2267006005, 2267006010, 2267006015, 2267006025, 2267006030	Mobile Sources; LPG; Commercial Equipment	Nonroad	Commercial + industrial land	Commercial equipment in surrogate file uses commercial + industrial land
2268002081	Mobile Sources; CNG; Construction and Mining Equipment; Other Construction Equipment	Nonroad	Inverse pop. density using total tract area	Construction and mining equipment in surrogate file use inverse pop. density
2268003020, 2268003030, 2268003040, 2268003060, 2268003070	Mobile Sources; CNG; Industrial Equipment	Nonroad	Industrial land	Industrial equipment in surrogate files uses industrial land
2268005050, 2268005055, 2268005060	Mobile Sources; CNG; Agricultural Equipment	Nonroad	farmland + orchard land	Ag equipment in surrogate file uses farmland + orchard land
2268006005, 2268006010, 2268006015, 2268006020	Mobile Sources; CNG; Commercial Equipment	Nonroad	Commercial + industrial land	Commercial equipment in surrogate file uses commercial + industrial land
2268010010	Mobile Sources; CNG; Industrial Equipment; Other Oil Field Equipment	Nonroad	Industrial land	Industrial equipment in surrogate files uses industrial land

Table B-2. Continued.

SCC code(s)	Description	Inventory	Surrogate	Reasons for surrogate choice
2270002006	Mobile Sources; Off-highway Vehicle Diesel; Construction and Mining Equipment; Tampers/Rammers	Nonroad	Inverse pop. density using total tract area	Construction and mining equipment in surrogate file use inverse pop. density
2280002100, 2280002200, 2280003100, 2280003200	Mobile Sources; Marine Vessels	Nonroad	Water	Other marine vessel SCC codes in surrogate file use water
2285002006, 2285002007, 2285002008,	Mobile Sources; Railroad Equipment	Nonroad	Railway miles	Other railroad SCC codes in surrogate file use railway miles

Table B-3. Surrogates to assign to SCC codes in 2000, 2010, and 2020 inventories.

SCC code(s)	Description	Inventory	Surrogate	Reasons for surrogate choice
10200401, 10200601	External Combustion Boilers; Industrial	Nonpoint	Industrial land	Other industrial boiler SCC codes use industrial land
20100102, 20100201, 20100202	Internal Combustion Engines; Electric Generation	Nonpoint	Industrial land	In 1996 NATA, 20100202 uses industrial land
2302002100, 2302002200	Industrial Processes; Food and Kindred Products: SIC 20; Commercial Charbroiling	Nonpoint	Food, drug, and chemical industry	2302000000, a related SCC codes uses this surrogate
2310001000	Industrial Processes; Oil and Gas Production: SIC 13	Nonpoint	Metals and minerals industry	2310000000, a related SCC code uses this surrogate
2501055120	Storage and Transport; Petroleum and Petroleum Product Storage	Nonpoint	Refineries and tank farms	2501050000, a related SCC code uses this surrogate
2505040120	Storage and Transport; Petroleum and Petroleum Product Transport; Pipeline; Gasoline	Nonpoint	Rural land	
49000199	Petroleum and Solvent Evaporation; Organic Solvent Evaporation; Solvent Extraction Process; Other Not Classified	Nonpoint	Heavy Industrial	Industrial process
50300601	Waste Disposal; Solid Waste Disposal - Industrial; Landfill Dump; Waste Gas Flares	Nonpoint	Low intensity residential land	2620000000, a landfill SCC code, uses this surrogate
2270002006	Mobile Sources; Off-highway Vehicle Diesel; Construction and Mining Equipment; Tampers/Rammers	Nonroad	Housing change and population	2270002XXX SCC codes use this surrogate

Table B-3. Continued.

SCC code(s)	Description	Inventory	SCC code to mimic	Description	Reasons for choosing SCC
30107101	Industrial Processes; Chemical Manufacturing; Hydrogen; Reformers	Point	30107001	Industrial Processes; Chemical Manufacturing; Inorganic Chemical Manufacturing (General); Fugitive Leaks	Related SCC code
30112558	Industrial Processes; Chemical Manufacturing; Chlorine Derivatives; Chloromethanes via MH & MCC Processes: Chloroform Condenser	Point	30112555	Industrial Processes; Chemical Manufacturing; Chlorine Derivatives; Vinylidene Chloride: Fugitive Emissions	Related SCC codes use same profiles
30130115	Industrial Processes; Chemical Manufacturing; Chlorobenzene; Atmospheric Distillation Vents	Point	30130110	Industrial Processes; Chemical Manufacturing; Chlorobenzene; Catalyst Incineration	Related SCC codes use same profiles
30301582	Industrial Processes; Primary Metal Production; Integrated Iron and Steel Manufacturing (See also 3-03-008 & 3-03-009); Miscellaneous Combustion Sources: Boilers	Point	30301499	Industrial Processes; Primary Metal Production; Barium Ore Processing; Other Not Classified	Similar SCC codes use same profiles
30600518	Industrial Processes; Petroleum Industry; Wastewater Treatment; Petroleum Refinery Wastewater System: Weir	Point	30600517	Industrial Processes; Petroleum Industry; Wastewater Treatment; Petroleum Refinery Wastewater System: Non-aerated Impoundment	Related SCC codes use same profiles
30601601, 30601604	Industrial Processes; Petroleum Industry; Catalytic Reforming Unit	Point	2306000000	Industrial Processes; Petroleum Refining: SIC 29; All Processes; Total	2306000000 is most general of refining SCC codes. Using the profile associated with 2306000000 will yield a flat profile, all hours the same.

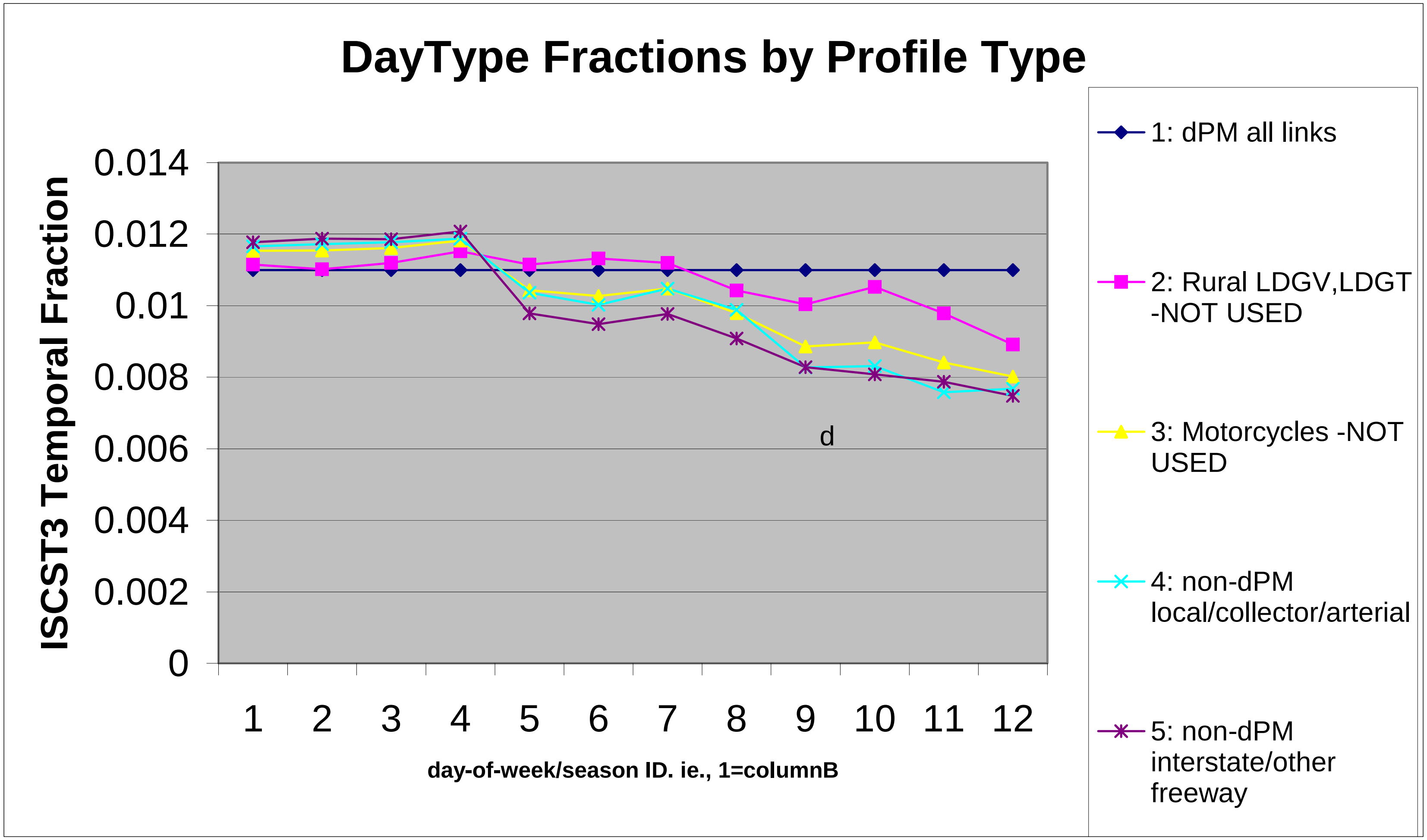


Figure B-1. Onroad temporal profiles for onroad link emissions. Adapted from Figure C-1EMS-HAP User's Guide, Appendix C.

Table B-4. Onroad gasoline SCC codes mapped to the GAS_INTRST SCC code.

SCC code	Description	SCC code	Description
2201001110	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Interstate: Rural Total	2201001230	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Interstate: Urban Total
2201001250	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Other Freeways and Expressways: Urban Total	2201020110	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Interstate: Rural Total
2201020230	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Interstate: Urban Total	2201020250	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Other Freeways and Expressways: Urban Total
2201040110	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Interstate: Rural Total	2201040230	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Interstate: Urban Total
2201040250	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Other Freeways and Expressways: Urban Total	2201070110	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Interstate: Rural Total
2201070230	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Interstate: Urban Total	2201070250	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Other Freeways and Expressways: Urban Total
2201080110	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Interstate: Rural Total	2201080230	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Interstate: Urban Total
2201080250	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Other Freeways and Expressways: Urban Total		

Table B-5. Onroad gasoline SCC codes mapped to the GAS_LOCAL SCC code.

SCC code	Description	SCC code	Description
2201001130	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Other Principal Arterial: Rural Total	2201001150	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Minor Arterial: Rural Total
2201001170	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Major Collector: Rural Total	2201001190	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Minor Collector: Rural Total
2201001210	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Local: Rural Total	2201001270	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Other Principal Arterial: Urban Total
2201001290	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Minor Arterial: Urban Total	2201001310	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Collector: Urban Total
2201001330	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Vehicles (LDGV); Local: Urban Total	2201020130	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Other Principal Arterial: Rural Total
2201020150	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Minor Arterial: Rural Total	2201020170	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Major Collector: Rural Total
2201020190	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Minor Collector: Rural Total	2201020210	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Local: Rural Total
2201020270	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Other Principal Arterial: Urban Total	2201020290	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Minor Arterial: Urban Total
2201020310	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Collector: Urban Total	2201020330	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 1 (LDGT1); Local: Urban Total
2201040130	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Other Principal Arterial: Rural Total	2201040150	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Minor Arterial: Rural Total
2201040170	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Major Collector: Rural Total	2201040190	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Minor Collector: Rural Total
2201040210	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Local: Rural Total	2201040270	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Other Principal Arterial: Urban Total
2201040290	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Minor Arterial: Urban Total	2201040310	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Collector: Urban Total
2201040330	Mobile Sources; Highway Vehicles - Gasoline; Light Duty Gasoline Trucks 2 (LDGT2); Local: Urban Total	2201070130	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Other Principal Arterial: Urban Total

Table B-5. Continued.

SCC code	Description	SCC code	Description
2201070150	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Minor Arterial: Rural Total	2201070170	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Major Collector: Rural Total
2201070190	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Minor Collector: Rural Total	2201070210	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Local: Rural Total
2201070270	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Other Principal Arterial: Urban Total	2201070290	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Minor Arterial: Urban Total
2201070310	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Collector: Urban Total	2201070330	Mobile Sources; Highway Vehicles - Gasoline; Heavy Duty Gasoline Vehicles (HDGV); Local: Urban Total
2201080130	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Other Principal Arterial: Rural Total	2201080150	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Minor Arterial: Rural Total
2201080170	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Major Collector: Rural Total	2201080190	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Minor Collector: Rural Total
2201080210	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Local: Rural Total	2201080270	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Other Principal Arterial: Urban Total
2201080290	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Minor Arterial: Urban Total	2201080310	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Collector: Urban Total
2201080330	Mobile Sources; Highway Vehicles - Gasoline; Motorcycles (MC); Local: Urban Total		

Table B-6. Onroad diesel SCC codes mapped to the DIE_INTRST SCC code.

SCC code	Description	SCC code	Description
2230001110	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Interstate: Rural Total	2230001230	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Interstate: Urban Total
2230001250	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Other Freeways and Expressways: Urban Total	2230060110	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Interstate: Rural Total
2230060230	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Interstate: Urban Total	2230060250	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Other Freeways and Expressways: Urban Total
2230071110	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Interstate	2230071230	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Interstate
2230071250	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Other Freeways and expressways	2230072110	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Interstate
2230072230	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Interstate	2230072250	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Other Freeways and expressways
2230073110	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Interstate	2230073230	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Interstate
2230073250	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Other Freeways and expressways	2230074110	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Interstate
2230074230	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Interstate	2230074250	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Other Freeways and expressways
2230075110	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Interstate	2230075230	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Interstate
2230075250	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Other Freeways and expressways		

Table B-7. Onroad diesel SCC codes mapped to the DIE_LOCAL SCC code.

SCC code	Description	SCC code	Description
2230001130	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Other Principal Arterial: Rural Total	2230001150	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Minor Arterial: Rural Total
2230001170	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Major Collector: Rural Total	2230001190	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Minor Collector: Rural Total
2230001210	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Local: Rural Total	2230001270	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Other Principal Arterial: Urban Total
2230001290	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Minor Arterial: Urban Total	2230001310	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Collector: Urban Total
2230001330	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Vehicles (LDDV); Local: Urban Total	2230060130	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Other Principal Arterial: Rural Total
2230060150	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Minor Arterial: Rural Total	2230060170	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Major Collector: Rural Total
2230060190	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Minor Collector: Rural Total	2230060210	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Local: Rural Total
2230060270	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Other Principal Arterial: Urban Total	2230060290	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Minor Arterial: Urban Total
2230060310	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Collector: Urban Total	2230060330	Mobile Sources; Highway Vehicles - Diesel; Light Duty Diesel Trucks (LDDT); Local: Urban Total
2230071130	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Other Principal Arterial	2230071150	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Minor Arterial
2230071170	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Major Collector	2230071190	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Minor Collector
2230071210	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Rural Local	2230071270	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Other Principal Arterial
2230071290	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Minor Arterial	2230071310	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Collector
2230071330	Mobile Sources; Highway Vehicles - Diesel; 2B Heavy Duty Diesel Vehicles; Urban Local	2230072130	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Other Principal Arterial
2230072150	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Minor Arterial	2230072170	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Major Collector
2230072190	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Minor Collector	2230072210	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Rural Local
2230072270	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Other Principal Arterial	2230072290	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Minor Arterial

Table B-7. Continued.

SCC code	Description	SCC code	Description
2230072310	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Collector	2230072330	Mobile Sources; Highway Vehicles - Diesel; Light Heavy Duty Diesel Vehicles; Urban Local
2230073130	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Other Principal Arterial	2230073150	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Minor Arterial
2230073170	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Major Collector	2230073190	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Minor Collector
2230073210	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Rural Local	2230073270	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Other Principal Arterial
2230073290	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Minor Arterial	2230073310	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Collector
2230073330	Mobile Sources; Highway Vehicles - Diesel; Medium Heavy Duty Diesel Vehicles; Urban Local	2230074130	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Other Principal Arterial
2230074150	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Minor Arterial	2230074170	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Major Collector
2230074190	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Minor Collector	2230074210	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Rural Local
2230074270	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Other Principal Arterial	2230074290	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Minor Arterial
2230074310	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Collector	2230074330	Mobile Sources; Highway Vehicles - Diesel; Heavy Heavy Duty Diesel Vehicles; Urban Local
2230075130	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Other Principal Arterial	2230075150	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Minor Arterial
2230075170	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Major Collector	2230075190	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Minor Collector
2230075210	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Rural Local	2230075270	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Other Principal Arterial
2230075290	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Minor Arterial	2230075310	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Collector
2230075330	Mobile Sources; Highway Vehicles - Diesel; Buses Heavy Duty Diesel Vehicles; Urban Local		

APPENDIX C:

BENZENE HEALTH EFFECTS LITERATURE REVIEW MEMO

INDUSTRIAL ECONOMICS, INCORPORATED
2067 Massachusetts Avenue Cambridge, Massachusetts 02140 Telephone 617/354-0074 Facsimile 617/354-0463

MEMORANDUM

5 July, 2005

TO:	Jenny Craig, EPA/OPAR
CC:	Nona Smoke, EPA/OPAR
FROM:	Tyra Gettleman and Henry Roman, IEc
SUBJECT:	Benzene Health Effects Literature Review

INTRODUCTION

In June 2003, Industrial Economics, Incorporated (IEc) and EPA presented to the Health Effects Subcommittee (HES) of the Scientific Advisory Board Council (SAB) an analytical plan for a case study estimating the health benefits of benzene reductions in the Houston area under the Clean Air Act Amendments (CAAA) of 1990. The analytical plan proposed to quantitatively estimate avoided cases of leukemia (all types) using a life-table approach that would allow us to assess the effects of changes in benzene exposures over time, and to implement a lag for the realization of benefits. The proposed life table approach would use risk estimates for leukemia from an analysis of an occupational cohort by Crump (1994), and would assume a five-year lag. The plan also proposed to semi-quantitatively assess changes in risk of decreased white blood cell counts by estimating changes in the numbers of individuals exposed above EPA's reference concentration (RfC) for benzene, and to qualitatively discuss other health endpoints (e.g., non-Hodgkin's lymphoma) that have been associated with benzene exposure in the literature.

The SAB HES in its response letter (EPA, 2004) made several recommendations concerning the analytical plan. They suggested that EPA take a closer look at studies of a large Chinese worker cohort exposed to benzene as a possible replacement for the risk estimates of Crump, which are based on a smaller cohort with fewer cases of leukemia. They also recommended that EPA consider studies of this larger cohort that suggest a non-linear concentration-response function for leukemia. Finally, they suggested that the proposed lag of five years did not make full use of available information, and recommended that EPA consider revising its approach to the lag issue after reviewing available epidemiological data.

IEc has conducted a literature review of the health effects of benzene to explore whether the analytical plan should be revised, either in response to the studies cited by the SAB, or because recent literature suggests additional health endpoints for us to consider in the Houston case study. The literature search thus focused on identifying evidence of non-leukemia health effects, defining the leukemia/benzene dose-response function, and characterizing the lag between benzene exposure and onset of leukemia. This review is not intended to replace EPA's evaluation of the literature on benzene health effects that was developed to support the benzene Integrated Risk Information System (IRIS) profile, but rather to complement that review with more recent data that may assist EPA in refining its analytical plan for the benzene case study.

LITERATURE SEARCH APPROACH

We conducted a search of peer-reviewed literature published in the past ten years pertinent to the health benefits portion of the benzene case study analytical plan. We identified relevant studies using the Dialog search engine. We began with a broad search of studies of the health effects of benzene, including leukemia, and then added keyword terms in two subsequent searches to focus on the dose-response relationship between benzene and leukemia and the latency period for developing leukemia. In addition, we conducted more focused searches, using key words for specific health effects to ensure completeness.[1] We reviewed abstracts for those studies that we felt might be relevant to the literature review, based on the title. We then included all studies that we felt would provide valuable information on one of the three subject areas of the literature review, which included 46 studies.

RESULTS OF THE LITERATURE REVIEW

In this section, we present the results of our literature search grouped into three topic areas: evidence for specific health effects associated with benzene exposures; issues concerning the dose-response function for benzene-induced leukemia; and evidence of a lag period for the full realization of benefits following reductions in benzene exposure ("cessation lag"). In the first two sections, we both briefly review the conclusions that EPA reached in quantifying risk estimates for inhaled exposures of benzene in support of the IRIS profile, and we discuss the findings of additional studies uncovered during our literature search. In the third section, we discuss how latency estimates and other data from existing epidemiological studies help define the concept of cessation lag.

The literature review that we conducted focused on human studies. We reviewed 46 studies, including 21 cohort analyses, 12 case-control analyses, 8 reviews, 3 meta-analyses, 1

---

[1] We used the following key words: "benzene" and "leukemia" and ("latency" or "lag time" or "incubation period") and "dose response" in various combinations. The more specific searches included "benzene" and the following key words: ("hematologic" or "blood"); "Chinese worker"; "non-Hodgkin's lymphoma"; "Hodgkin's" "disease or lymphoma"; "myelodysplastic syndrome"; "multiple myeloma"; and "lymphohematopoeitic."

cross-sectional study, and one exposure validation study. Exhibit 1 below displays studies we identified in our search, grouped by study type.

EXHIBIT 1
LITERATURE SEARCH RESULTS

Cancer		
Study Type	***Citation***	***Endpoint(s)***
Case-Control	Finkelstein (2000)	Leukemia
Case-Control	Glass et al. (2003)	Leukemia, ANLL, CLL, CML, NHL, MM
Case-Control	Guenel et al. (2002)	Leukemia
Case-Control	Rushton and Romaniuk (1997)	Leukemia, AMML, CLL, ALL, CML
Case-Control	Schnatter et al. (1996a)	Leukemia, MM
Cohort	Adegoke et al. (2003)	Leukemia, ALL, AML, CML
Cohort	Bloemen et al. (2004)	Leukemia, CLL, ANLL, MM, NHL, HL
Cohort	Collins et al. (2003)	Leukemia, ANLL, CML, MM, NHL, HL
Cohort	Costantini et al. (2003)	Leukemia
Cohort	Crump (1994 & 1996)	Leukemia, AMML
Cohort	Hayes et al. (1997 & 2000)	Leukemia, ANLL, AML, CML, ALL, NHL
Cohort	Ireland et al. (1997)	Leukemia, MM, ANLL
Cohort	Paxton et al. (1987)	Leukemia
Cohort	Rinsky et al. (1981, 1987 & 2002)	Leukemia, MM, NHL
Cohort	Schnatter et al. (1996b)	Leukemia, AMML
Cohort	Silver et al. (2002)	Leukemia
Cohort	Sorahan et al. (2005)	Leukemia, AML, CML, CLL, NHL, HL
Cohort	Swaen et al. (2005)	Leukemia, MM, HL
Cohort	Wong (1995)	AMML, MM
Cohort	Yin et al. (1987 & 1996)	Leukemia, AML, CML, ALL, MM, NHL
Cohort/Case-Control	Rothman et al. (1997)	ANLL/MDS, Enzymatic genotypes
Exposure Validation	Dosemeci et al. (1996)	Validates exposure estimates in Chinese Worker Cohort
Meta-Analysis	Lamm et al. (2005)	NHL
Meta-Analysis	Sonoda et al. (2001)	MM
Meta-Analysis	Wong and Raabe (2000)	NHL
Review	Bergsagel et al. (1999)	MM
Review	Bezabeh et al. (1996)	MM
Review	Budinsky et al. (1999)	Exposure estimates in Chinese Worker Cohort
Review	Hayes et al. (2001)	Exposure estimates in Chinese Worker Cohort
Review	Savitz and Andrews (1997)	Leukemia and subtypes
Review	Utterback and Rinsky (1995)	Exposure estimates in Pliofilm Cohort
Review	Wong (1999 & 2002)	Exposure estimates in Chinese Worker Cohort
Non-Cancer		
Case-Control	Lan et al. (2004)	Decreased lymphocytes
Case-Control	Qu et al. (2002)	Decreased RBCs, WBCs, lymphocytes and neutrophils
Case-Control	Rothman et al. (1996a)	Decreased lymphocyte counts, benzene metabolites
Case-Control	Rothman et al. (1996b)	Decreased lymphocyte count, chromosome damage
Cross-Sectional	Collins et al. (1997)	Decreased lymphocytes
Biomarkers of Exposure		
Case-Control	Rappaport et al. (2002)	Benzene metabolites
Case-Control	Rothman et al. (1995)	Chromosome damage
Case-Control	Rothman et al. (1998)	Benzene metabolites

Acronyms: AMML = acute myelogenous and monocytic leukemia; ANLL = acute non-lymphocytic leukemia; ALL = acute lymphocytic leukemia; CLL = chronic lymphocytic leukemia; CML = chronic myelogenous leukemia; HL = Hodgkin's lymphoma; MDS = myelodysplastic syndrome; MM = multiple myeloma; NHL = non-Hodgkin's lymphoma; RBC = red blood cell; WBC = white blood cell.

Benzene Health Effects

This section describes the various health effects that we identified in the literature review as having a potential link to benzene exposures. A number of effects have been studied, with varying levels of support in the literature. Exhibit 2 presents IEc's assessment of the strength of evidence supporting a relationship to benzene exposure for each health effect.

EXHIBIT 2
BENZENE HEALTH EFFECTS WEIGHT OF EVIDENCE

Health Effect	***Strength of Evidence***
Leukemia (all types)	High
Acute Myelogenous	Medium
Acute Lymphocytic	Low
Chronic Myelogenous	Low
Chronic Lymphocytic	Low
Multiple Myeloma	Low
Myelodysplastic Syndrome	Low
Hodgkin's Lymphoma	Low
Non-Hodgkin's Lymphoma	Low
Decreased Lymphocytes	High

Leukemia

Significantly increased risks of leukemia have been consistently reported in benzene-exposed workers of various industries, leading EPA to classify inhaled benzene as a "known/likely" human carcinogen under the proposed 1996 cancer guidelines. In the EPA document *Carcinogenic Effects of Benzene: An Update* (EPA, 1998), it states "[e]pidemiologic studies and case studies provide clear evidence of a causal association between exposure to benzene and leukemia" (page 4). Our literature review also supports a link between benzene exposure and leukemia.

There are two cohorts in particular that EPA describes, which have been extensively studied and peer-reviewed. The first consists of a group of 1,717 white male workers employed in Pliofilm manufacturing plants located in Ohio between 1940 and 1972 (hereafter, the "Pliofilm Cohort").[2] The second is a cohort of 74,828 workers in a variety of industries in China employed between 1972 and 1987 (hereafter, the "Chinese Worker Cohort") studied jointly by the US National Cancer Institute (NCI) and the Chinese Academy of Preventative Medicine (CAPM). Results from these retrospective cohort studies indicate an association between exposure to a range of benzene concentrations and an elevated risk of leukemia (all types). Recent analyses comparing exposed workers to unexposed workers in the Chinese Worker

---

[2] Pliofilm is a glossy membrane made from rubber hydrochloride and used chiefly for water-resistant materials and packaging (Crump, 1994).

Cohort show significant elevated relative risks (RRs) of leukemia incidence of 2.6 (95%CI: 1.3, 5.7) (Yin et al., 1996) and 2.5 (95% CI: 1.2, 5.1) (Hayes et al., 1997). In other words, the exposed workers were roughly two and a half times more likely to develop leukemia than the unexposed workers. Similarly, a recent Pliofilm Cohort analysis found an elevated standardized mortality ratio (SMR) of 2.9 (no 95% CI provided), comparing the observed cases of leukemia in the cohort to an expected number of cases based on US sex- and age-specific rates.

Through our literature review, we identified several other recently published epidemiologic studies that have found an overall increase in risk of leukemia (all types) with exposure to benzene, or a trend of increasing relative risks with increased exposure to benzene (Ireland et al., 1997; Costantini et al., 2003; Adegoke et al., 2003; Sorahan et al., 2005; Guenel et al., 2002; Bloemen et al., 2004; Glass et al., 2003; Collins et al., 2003).

Leukemia Subtypes

There are four types of leukemia: Acute Myelogenous Leukemia (AML) (also referred to as Acute Myelogenous and Monocytic (AMML) or Acute Non-Lymphocytic Leukemia (ANLL)), Acute Lymphocytic Leukemia (ALL), Chronic Myelogenous Leukemia (CML), and Chronic Lymphocytic Leukemia (CLL). The strength of evidence supporting a link between benzene and specific types of leukemia varies. AML has the most evidentiary support for a link with benzene exposures out of all of the four subtypes of leukemia, but some of this evidence is conflicting. EPA concludes "[a] number of studies, including the Pliofilm cohort, have indicated that benzene exposure is associated with various types of lymphohematopoietic neoplasia other than ANLL (Savitz and Andrews, 1996). However, the specific types associated with benzene exposure remain unidentified" (EPA, 1998, page 5).

Our research uncovered associations between benzene and AML in the literature, including both of the major cohort studies. The Chinese Worker Cohort found an elevated RR of ANLL incidence of 3.0 (95% CI: 1.0, 8.9) and 3.1 (95% CI: 1.2, 10.7) (Hayes et al., 1997; Yin et al., 1996) and the Pliofilm Cohort identified a RR of AML deaths of 5.03 (95% CI: 1.84, 10.97) (Wong, 1995). The Pliofilm Cohort analysis also found evidence for an increasing trend of AML with increasing cumulative exposure to benzene (Crump, 1994, 1996; Wong, 1995). In addition, a study by Glass et al. (2003) found a significantly increased relative risk of ANLL among petroleum workers at much lower levels of exposures. The authors found a RR of 7.17 (95% CI: 1.27, 40.4) for workers exposed to greater than 8 ppm-years of benzene compared with those exposed to less than or equal to 4 ppm-years. The wide confidence bounds associated with this estimate, however, are evidence of statistical instability, calling into question the validity of the Glass et al. results. Other recent studies that we identified through the literature search have not found the same strength of association, finding only non-significantly elevated risks of AML with benzene exposure (Rushton and Romaniuk, 1997; Ireland et al., 1997; Adegoke et al., 2003; Sorahan et al., 2005; Bloemen et al., 2004; Guenel et al., 2002; Collins et al., 2003). These studies suffer from methodological weaknesses such as small numbers of cases and possible exposure misclassification, which may have limited their ability to detect an association.

Very few studies have shown an increase in risk due to the other leukemia subtypes aside from AML. EPA (1998) concluded that there may be evidence supporting an association of benzene with CML and CLL. They cite a study by Rushton and Romaniuk (1997) that found a non-significant increase in risk of CLL in petroleum workers in the UK whose benzene exposure increased with duration of employment.

We attempted, through the literature search, to find evidence supporting a link between benzene and specific non-AML leukemia subtypes. We found that Hayes et al. (2000) reported non-significant elevated relative risks for CML (RR = 2.6) and ALL (RR = 2.8), but also reported small numbers of cases for these two subtypes, making the results unstable. Another recent study found significant results for CML with an odds ratio (OR) of 2.4 (95%CI: 1.3, 4.7) comparing workers that were ever exposed with those who were never exposed to benzene. In addition, the authors found a significant trend for risk of CML with increasing duration of exposure (Adegoke et al., 2003). However, this study used self-reported exposure estimates, which are likely to be affected by recall bias, so these results should be interpreted with caution.[3] Several studies found no significant results for the non-AML subtypes (Sorahan et al., 2005; Bloemen et al., 2004; Ireland et al., 1997; Glass et al., 2003; Collins et al., 2003). Because chronic leukemias are rare, and because ALL tends to occur in children more often than adults, it is possible that the occupational cohort studies available do not have large enough study populations to detect associations between benzene and these leukemia subtypes, especially if the association is weak. Furthermore, with such small numbers of cases, any errors in disease classification due to imprecise or inaccurate diagnoses could have a substantial impact on whether or not a study finds an association.

Hodgkin's and Non-Hodgkin's Lymphomas

Few studies exist that examine an association between benzene exposure and either Hodgkin's Lymphoma (HL) or Non-Hodgkin's Lymphoma (NHL).[4] The IRIS support document for benzene carcinogenicity cites results from the Chinese Worker Cohort that showed a significantly elevated relative risk of developing NHL for benzene workers with 10 or more years of benzene exposure (RR = 4.2 (95%CI: 1.1, 15.9) (Hayes et al., 1997). However, this estimate is fairly unstable, as indicated by the wide confidence bounds, and has not been confirmed through the results of other, more recent epidemiologic studies (Sorahan et al., 2005; Bloemen et al., 2004; Schnatter et al., 1996a; Glass et al., 2003; Collins et al., 2003). In addition, two meta-analyses, one of 26 cohorts of petroleum workers, which included a total of 506 deaths from NHL (Wong and Raabe, 2000), and one consisting of 21 occupational study groups and 404 cases of NHL (Lamm et al., 2005) did not find positive associations with exposure to benzene, reporting SMRs of 0.90 (95%CI: 0.82, 0.98) and 1.04 (95%CI: 0.94, 1.14) respectively. An abstract presented at the *Recent Advances in Benzene Toxicity* conference in Munich, Germany, reviewed the relationship between benzene and NHL, and concluded that "most studies do not find an association between benzene exposure and NHL" (Lamm et al., 2004).

---

[3] Recall bias occurs when cases and controls differentially recall events related to their exposure. This can occur because cases tend to scrutinize their exposure history more closely than controls (Gerstman, 1998).

[4] Hodgkin's Lymphoma and Non-Hodgkin's Lymphoma are both cancers that start in the lymphatic tissue, often in the lymph nodes. Cancerous cells in Hodgkin's disease are called Reed-Sternberg (R-S) cells, and are different from the cells of non-Hodgkin's lymphoma. Scientists believe that the R-S cells are a type of malignant B lymphocyte (Medline Plus, a service of the US National Library of Medicine: http://medlineplus.gov/).

Of the studies that we identified that looked at the risk of HL related to benzene exposure, none of them found positive results (Sorahan et al., 2005; Bloemen et al., 2004; Ireland et al., 1997; Schnatter et al., 1996a; Collins et al., 2003; Swaen et al., 2005).

Multiple Myeloma

A few studies cited in the IRIS support document, including a case study (DeCoufle et al., 1983) and the analyses of Pliofilm Cohort (Rinsky et al., 1987; Wong, 1995) found an increased risk of multiple myeloma associated with benzene exposure. Several recent studies, however, including large-scale cohort studies, have failed to confirm this, and have found no associations or weak associations between benzene and multiple myeloma (Hayes et al., 1997; Ireland et al., 1997; Sorahan et al., 2005; Schnatter et al., 1996a; Glass et al., 2003; Swaen et al., 2005; Collins et al., 2003). In addition, two reviews examining the literature linking multiple myeloma and benzene exposure conclude that "benzene exposure is not a likely causal factor for multiple myeloma" (Bezabeh et al., 1996) and that there is "no scientific evidence to support a causal relationship between exposure to benzene … and the risk of developing multiple myeloma" (Bergsagel et al., 1999). A meta-analysis of case-control studies supports these conclusions, finding an OR of 0.74 (95%CI: 0.6, 0.9) for multiple myeloma for those with occupational exposure to benzene or organic solvents (Sonoda et al., 2001).

Myelodysplastic Syndrome

We found no evidence for an association between benzene and myleodysplastic syndrome (MDS) alone, but found a positive RR for a combined outcome of MDS/ANLL in the Chinese Worker Cohort of 4.1 (95%CI: 1.4, 11.6) (Hayes et al., 1997). The fact that MDS is a known precursor to AML makes it difficult to assess the effects of benzene on MDS separately from those on AML. In addition, Hayes et al. (2001) notes that Chinese Workers diagnosed with MDS were originally diagnosed as having ANLL. The similarity in clinical characteristics of these two conditions could lead to misclassification of the outcome, making an analysis of the effects of benzene on MDS challenging.

Additional Cancerous Endpoints

EPA discusses other cancerous endpoints in addition to leukemia in their benzene carcinogenicity update (EPA, 1998). They cite animal studies that have found cancer in multiple target organ sites such as oral and nasal cavities, liver, forestomach, preputial gland, lung, ovary, and mammary gland. We found no epidemiologic evidence to support these associations in our literature review.

In addition, EPA's carcinogenic assessment of benzene discusses the evidence for a link between parental occupational exposure to benzene and childhood leukemia. Although a handful of studies have found positive associations there is not conclusive evidence for this link (see EPA, 1998, page 42). EPA concludes "data to make quantitative adjustments for [increased risk due to parental occupational exposures to benzene] do not exist at this time" (EPA, 1998, page 42). We did not find any additional studies on this topic in our literature search.

Non-Cancerous Effects

Benzene has been associated with a number of non-cancer health effects; however, many of these appear unlikely to occur at levels expected to be found in ambient air (less than 10 ppb, based on EPA's NATA study). Benzene exposure at high concentrations has been associated with various hematological abnormalities, including aplastic anemia.

EPA developed a reference concentration (RfC) of 0.03 mg/m^3, based on a critical effect of decreased lymphocyte count from a cross-sectional study by Rothman et al. (1996a), which analyzed 44 members of the Chinese Worker Cohort data. This study found blood cell effects at exposure concentrations of about 8 ppm. The EPA support document for non-cancerous effects (EPA, 2002) identified additional studies that have also found decreases in hematologic factors (Ward et al., 1996; Bogadi-Sare et al., 2000) but that do not provide sufficient data to assess a LOAEL or NOAEL. In addition, EPA recognized some studies that did not find positive associations between benzene and hematologic factors, making these studies unsuitable for establishing a LOAEL (Khuder et al., 1999; Collins et al., 1991).

We identified two recent case-control studies that found statistically significant decreases in lymphocyte counts in workers with low exposure concentrations of less than 1 ppm (Lan et al., 2004) and less than 0.25 ppm (Qu et al., 2002). These studies both used high quality exposure assessment (personal monitors) and controlled for important confounding factors. In addition, we found a study by Collins et al. (1997) that does not show positive results for workers with an average exposure of 0.55 ppm, but this study relied on historical exposure data and blood samples collected through a medical surveillance program, making the results somewhat uncertain.

Issues Related to the Leukemia/Benzene Dose-Response Function

Epidemiologic Evidence for the Dose-Response Function

EPA supports the use of data from the Pliofilm cohort for quantifying the dose-response relationship between inhaled benzene and leukemia. A range of unit risk values is provided on EPA's IRIS for an individual exposed over a lifetime to 1 µg/m^3 of benzene in air. The unit risk range, 2.2×10^{-6} to 7.8×10^{-6}, is based on Crump's 1994 analysis of the Pliofilm cohort, with lower and upper bounds derived using a linear dose-response model and Paustenbach (1992) and Crump and Allen (1984) exposure estimates, respectively.

EPA recommends using the Pliofilm cohort because of methodological weaknesses in the Chinese Worker Cohort, such as confounding by exposure to other chemicals, and potential exposure misclassification. EPA states in the IRIS support document for benzene that "[t]he derivation of the cohort from many different factories across China suggested the possibility that this cohort was exposed to mixtures of many different chemicals…[which] could have produced confounding effects, especially if exposures were to chemicals that increase the risk of leukemia" (EPA, 1998, page 13). The Pliofilm cohort workers, on the other hand, were exposed primarily to benzene, with little exposure to other chemicals. EPA also found that the exposure assessment used with the Chinese Worker Cohort was flawed. EPA states that "only 38% of the exposure estimates were based upon actual measurements of benzene concentrations; the remainder were numbers generated by factory industrial hygienists based upon their estimates of benzene concentrations" (EPA, 1998, page 13). Therefore, EPA concludes that the dose per individual could have been subject to random error and to bias, which could have affected the shape of the dose-response relationship.

We found several additional cohort and case-control studies examining the relationship between benzene exposure and leukemia (Guenel et al., 2002; Costantini et al., 2003; Adegoke et al., 2003; Sorahan et al., 2005; Bloemen et al., 2004; Rushton and Romanieuk, 1997; Schnatter et al., 1996a; Swaen et al., 2005; Collins et al., 2003; Glass et al., 2003). The SAB HES, in their review of our original analytical plan, cited two of these studies (Rushton and Romanieuk (1997) and Schnatter (1996a)) as examples of studies finding an association at levels closer to those likely to be modeled in the case study. These studies involve analyses of two cohorts of petroleum workers, one in the United Kingdom and one in Canada, who are known to have low average exposures (e.g., less than 5 ppm (Rushton and Romanieuk, 1997)). In a nested case-control analysis, Rushton and Romanieuk compared 91 cases of leukemia to matched controls and found a slightly elevated relative risk for increasing cumulative exposure (1.004 (95% CI: 0.99, 1.02). However, incomplete or missing exposure information limit the usefulness of these results. Twenty percent of work histories were incomplete, and assumptions were made for missing exposure data such as hygiene data for base estimates, data on closed terminals, and product source, which contributed to uncertainties in the exposure estimates (Rushton and Romanieuk, 1997). Similar results were found for the Canadian cohort, which compared 14 cases of leukemia with matched controls and found a non-significant odds ratio of 1.002 for each ppm-year of exposure (95% CI: 0.989, 1.015). The authors of this study acknowledge that the lack of finding of a dose-response relationship between cumulative benzene exposure at low levels and leukemia may be due to limited statistical power deriving from small sample size.

The other recent cohort and case-control studies that have looked at the association between benzene and leukemia suffer from methodological weaknesses, such as small cohort size, insufficient exposure assessment, and potential confounding of other exposures that limit the usefulness of these studies for our analysis (see Attachment 1 for a summary of the limitations of each study). We will focus the remainder of this discussion on the most extensively studied and peer-reviewed cohorts; the Pliofilm Cohort and the Chinese Worker Cohort. Exhibit 3 below compares the characteristics of the two cohorts, highlighting methodological strengths and weaknesses of each.

EXHIBIT 3
COMPARISON OF THE PLIOFILM COHORT AND THE CHINESE WORKER COHORT

	Pliofilm Cohort	**Chinese Worker Cohort**
Description of Industry	Workers in Pliofilm manufacturing plants in two locations in Ohio	Workers in 672 factories in 12 cities of China employed in a number of industries such as painting, printing, footwear, rubber, and chemical
Cohort Size/Number of Leukemia Cases	1,717 white males/14 cases of leukemia	74,828 benzene exposed workers/47 cases of leukemia
Dates of Employment/ Follow-up	1939-1976/Follow-up through 1987 Rinsky et al. (2002) followed subjects through 1996.	1972-1987
Exposure Levels with Positive Effects	>40 ppm (cumulative exposure)	<10 ppm (average exposure); <40 ppm-years (cumulative exposure)
Exposure Assessment Method	Crump and Allen (1984) updated the exposure assessments made in Rinsky et al. (1981) by estimating calendar-specific benzene concentrations for various work areas, allowing for the creation of a complete exposure profile for each worker. Paustenbach et al. (1992) made a detailed reevaluation of exposures in this cohort that incorporated information obtained from historical records and interviews with former workers. This newer assessment accounted for dermal exposures, short-term high-level exposures, respirator use, biases of sampling devices used in earlier years, and a previously unaccounted for shutdown of the St. Mary's plant during World War II.	Work history data for each worker was merged with exposure data based on job title, using measurement data and historical information such as product use in each factory (Dosemeci et al., 1994).
Major Results	RR comparing total observed leukemia deaths to expected deaths, based on US sex- and age-specific rates = 2.9. Found that multiplicative, linear models were the best fit for the dose-response data (Crump, 1994).	Incidence of leukemia in all exposed subjects compared to unexposed subjects, RR = 2.5 (1.2, 5.1), controlling for age and sex. Significant trend for increasing RRs with increasing exposure category (p = 0.04) (Hayes et al., 1997).
Strengths	-Workers exposed to benzene primarily (not likely to have significant exposures to other carcinogens) -Thorough exposure assessment -Dose-response relationship investigated for leukemia deaths, and betas reported per ppm-year	-Larger number of cases of leukemia -Positive results seen at lower benzene exposures
Limitations	-Relatively smaller number of cases of leukemia -Benzene exposures higher than those experienced by the general public -No measurement data available prior to 1946.	-Workers may have been exposed to a variety of other carcinogens in addition to benzene -Exposure estimates have been criticized as underestimates (only 38/% of exposure estimates were based on measurements) -No results for increased risk per ppm-year

Health Endpoints Considered: One difference between the two cohort studies is the endpoint examined. The Chinese Worker Cohort reported leukemia incidence, while the Pliofilm Cohort reported leukemia deaths. Therefore, the Pliofilm study could have underestimated cases of leukemia by only reporting deaths. However, survival rates for leukemia during the time of the Pliofilm Cohort were low, leading us to assume that leukemia deaths and leukemia incidence may be considered reasonably equivalent.

Exposure Assessment: Both the Chinese Worker Cohort and the Pliofilm Cohort analyses are retrospective cohort studies, making historical exposure assessment challenging. Dosemeci et al. (1994) state that 38 percent of the exposure estimates in the Chinese Worker Cohort are based on monitoring data. The Pliofilm Cohort data are based on monitoring data that varies in quantity with time and by site. For instance, the number of samples increases over time, with very little data on exposures before 1950. Also, the Akron I plant has virtually no measurement data, while the St. Mary's plant has a great deal.[5] The inconsistency in monitoring data for both cohorts makes the exposure assessments for both of these analyses somewhat uncertain.

Exposure assessment for the Pliofilm Cohort has been investigated by three separate research groups, Rinsky et al. (1981 & 1987), Crump and Allen (1984), and Paustenbach et al. (1992), yielding a variety of results. The different exposure assessment results of these three analyses can be attributed to various assumptions made by the investigators in relation to exposure of the workers, such as exposure concentrations experienced before sufficient monitoring data was available. Paustenbach et al. estimates are the highest, followed by Crump and Allen, and then Rinsky et al. Accordingly, the Rinsky et al. estimates yield higher relative risks than the other two exposure estimates. The estimates by Paustenbach et al. (1992) have been criticized in a paper by Utterback and Rinsky (1995). These authors contend that the Paustenbach et al. exposure estimates were based upon worst-case assumptions for the exposure scenarios that existed during the early years of the cohort. In addition, Utterback and Rinsky noted that prolonged exposure to the high levels of benzene estimated by Paustenbach et al. would have resulted in much higher prevalence of benzene poisoning than was actually seen in the cohort. EPA points out, however, that despite differences in the three sets of exposure estimates, the cumulative SMRs from the three studies differ by no more than a factor of 2.5 (see EPA, 1998, Table 2, Page 10).

The Chinese Worker Cohort has one set of exposure estimates, as described by Dosemeci et al. (1994). These exposure estimates have been criticized by Wong (1999 & 2002) and Budinsky et al. (1999). The authors state that these exposure estimates are not consistent with exposure measurements provided by the CAPM investigators before NCI's involvement or with studies providing air monitoring data. Wong and Budinsky et al. conclude that Dosemeci et al. exposure estimates are likely to be underestimated, based on these other available measurements. Budinsky et al. also points out that benzene poisoning is a biomarker for benzene exposure, and incidence of chronic benzene poisoning seen in a study based on the Chinese Worker Cohort (Yin et al., 1987) suggests higher exposures were experienced by the workers than those reported in Dosemeci et al. (1994). The review authors also cite a number of other limitations of the

---

[5] See Paustenbach et al. (1992), Figure 3, page 183 for details on sampling data.

exposure assessment, such as poor exposure assumptions (relating to the percentage of benzene in various products used in the factories), inconsistencies in calculating individuals' exposures, wide exposure categories, and an inadequate validation study. Specifically, Wong (1999) states that the results of the validation study by Dosemeci et al. (1996) only indicate that exposure estimates are valid in relation to each other, and one could find a similar upward trend as described in the validation results if benzene exposure levels were underestimated.

Authors of the Chinese Worker Cohort analyses published a response to the criticisms outlined by Wong and Budinsky et al. (Hayes et al., 2001). They acknowledge that the estimates are not consistent with exposures in recently published papers using monitoring data, but argue that these measurements were only taken in a small number of workplaces that would not necessarily be reflective of concentrations found in all of the 672 factories in the Chinese Worker Cohort studies. In addition, they state that exposure measurements taken during the CAPM studies were not systematized, were taken during a time period when benzene exposures were higher, and were taken at a single point in time, making them less suitable for personal exposure assessment. They also defend their estimates against internal consistencies, saying that there were differences in reporting between two CAPM papers, but that these exposure estimates were not carried through to the NCI-CAPM studies. Finally, the authors support their validation study (Dosemeci et al., 1996) by stating that the results showed a clear dose-response relationship between benzene exposures and benzene poisoning, which provides evidence of the predictive capacity of the exposure assessment and of the accuracy in the quantitative estimation of benzene exposure.

Confounding Factors: The two main cohorts also differ in the amount of exposure that the workers had to other potential carcinogens. Wong (1999) mentions that in the original analysis by Yin et al. (1987), 95 percent of those in the Chinese Worker Cohort were exposed to chemicals other than benzene. He goes on to say that the control workers had no known exposure to benzene or other occupational carcinogens, meaning that increased risk in health effects seen in the exposed workers may reflect the effects of other occupational carcinogens in addition to benzene. The workers in the Pliofilm Cohort, on the other hand, were exposed primarily to benzene and it is likely that increased risks found in these analyses were due to benzene exclusively.

The Hayes et al. (2001) response states that the risks for ANLL/MDS were systematically increased across all of the diverse industries studied, which leads to the conclusion that the associations were due to the common exposure to benzene, rather than other carcinogens. Hayes also points out that other industrial exposures linked to benzene, such as ionizing radiation, butadiene, and ethylene oxide are unlikely to contribute to the associations seen based on occupational data from the cohort. Finally, Hayes states that elevated risk for ANLL were found in the painters who used benzene-containing paint but that painters not exposed to benzene do not show increased risks for leukemia.

Shape of the Dose-Response Function

The shape of the dose-response function for leukemia and benzene is uncertain, with different studies suggesting one or more possible functional forms (e.g., linear, supralinear). EPA acknowledges this uncertainty in the dose-response function due to questions about the mode of action for benzene-induced leukemia. They indicate that there is conflicting information on the possibility of a threshold in the dose-response function, as well as existing data suggesting a supralinear shape at low doses. EPA concludes that the lack of effects at low levels seen in some studies may not be indicative of a threshold, but instead may be due to lack of power in current data to examine low-dose effects of benzene. They also point out that if there are individual threshold levels, due to variability in sensitivity to benzene's effects, it is unlikely that a single threshold dose could apply to an entire population exposed to benzene. Lack of consistent statistical data, coupled with evidence from studies on the mode of action of benzene (e.g., studies on benzene metabolism and chromosomal damage), and high background levels of benzene in the environment[6] lead EPA to conclude that a linear dose-response function for benzene at low doses would be sufficiently conservative, stating that "there is insufficient evidence to reject this concept [of a linear extrapolation to low doses]" (EPA, 1998, page 37).

EPA notes in the IRIS support document for benzene (EPA, 1998) the existence of some evidence for a possible threshold of benzene exposure necessary to see increased risk of leukemia. We investigated the possibility of a threshold in our literature search. Schnatter et al. (1996b) reanalyzed the Pliofilm Cohort data, calculating average total concentration per person. The authors found a "critical" concentration of 35-40 ppm when a median exposure was used (using a combination of Rinsky (1981), Crump and Allen (1984), and Paustenbach (1992) exposure estimates). In addition, Pliofilm cohort data has not found significant increases for leukemia below 40 ppm-years of exposure, which suggests a potential threshold. However, all of these findings are uncertain due to low power of these studies at low levels of exposures.

EPA pointed out in the IRIS support document for benzene (EPA, 1998), that some evidence exists for a supralinear dose-response function. For instance, Hayes et al. (1997) found relative risks for leukemia that are significantly elevated at 10 ppm of benzene, but tend to plateau as the dose increases to higher levels. However, concerns about bias in the exposure assessment for the Chinese Worker Cohort data could have contributed to a spurious supralinear dose-response reported in the studies using the historical data to calculate cumulative exposure.[7]

Studies of benzene metabolism may give some insight into the shape of the dose-response function, since animal and human studies have shown that benzene metabolites may exert the carcinogenic effects of benzene (EPA, 1998). Rothman et al. (1996b) found that formation of urinary toxic metabolites decreased from 32 percent in workers exposed to <31 ppm of benzene in air to 24 percent in workers exposed above this level, suggesting that a plateau exists for benzene effects at higher exposures. Rothman et al. (1998) found that relative levels of the benzene metabolites hydroquinone and muconic acid decreased while phenol and catechol increased in the more highly exposed workers compared with the less exposed. The authors conclude that, assuming that hydroquinone is the toxic metabolite of benzene, the results suggest that "the risk for adverse health outcomes due to exposure to benzene may have a supralinear relation with external dose" (Rothman et al., 1998, page 711). The author does point out,

---

[6] High background levels of benzene could overwhelm defense mechanisms in the body that might otherwise show a threshold effect.

[7] The exposure estimates in the Chinese worker study are thought to be underestimated, which could lead to inflated relative risks at lower levels, producing a supralinear curve.

however, that urinary metabolites may not necessarily reflect metabolite concentrations in target tissues. Rothman et al. (1997) also investigated the hypothesis that individuals with mutations in genes affecting enzymes involved in benzene metabolism would be more susceptible to benzene poisoning. The authors' results supported this hypothesis, suggesting that there may be an enzyme-mediated process involved in benzene toxicity that could involve saturation of the enzyme at higher doses. This type of process would also support a supralinear dose-response curve. These studies, however, indicate that saturation in benzene concentrations likely occurs at levels much higher than those expected to be found in our case study.

We identified a study by Rappaport et al. (2002) investigating the presence of albumen adducts of benzene oxide (BO-Alb) and 1,4-benzoquionone (1,4-BQ-Alb) in the blood of workers exposed to low levels of benzene in China. Exposure in this study was measured with personal benzene monitors. The authors found a supralinear dose-response for benzene exposures and production of Bo-Alb and 1,4-BQ-Alb with deviations from linearity beginning at 1 ppm. The authors attributed this to saturable metabolism of benzene at that concentration.

We found further evidence for a supralinear dose-response function from another study by Rothman et al. (1995), which found an association between cumulative exposure to benzene and chromosome damage, which is thought to be a factor in the development of leukemia. The authors found a trend of increasing variants at a gene locus that suggests gene-duplicating mutations with increasing cumulative exposure to benzene. They reported a significant supralinear trend for this relationship ($p = 0.0002$). The results of this study may show a spurious supralinear dose-response relationship with benzene because of potential exposure misclassification, since its exposure estimates were based on Dosemeci et al. (1994).

Linear models were found to be the best fit in the Crump (1994) analysis of the Pliofilm Cohort. The author concluded that "[t]here was no indication of either [cumulative exposure]-dependent or intensity-dependent nonlinearity in the dose responses for any model based on the Crump and Allen exposure matrix" (Crump, 1994, page 234). Only borderline significant results were found for a intensity-dependent nonlinear model, using the Paustenbach exposure estimates.

EPA (1998) concludes that "[t]oo many questions remain about the mode of action for benzene-induced leukemia for the shape of the dose-response function to be known with certainty" (page 34). According to EPA's *Guidelines for Carcinogen Risk Assessment*, linear extrapolation to low doses should be used when there is insufficient data to establish a mode of action (MOA) as a default approach because linear extrapolation "generally is considered to be a health-protective approach" (EPA, 2005, page 3-21).

Cessation Lag

The term "cessation lag" refers to the estimate of how quickly cancer risks in a population will decline to a new steady-state level following a reduction in exposure. In their review of the analytical plan for the benzene case study, the SAB HES subcommittee suggested that we revisit our proposed five-year "cessation lag" for benzene-induced leukemias in light of evidence from available epidemiologic studies. Exhibit 4 summarizes the findings of the studies in our literature review with respect to the issue of latency or cessation lag of benzene-induced leukemias.

Only one study in Exhibit 4, Silver et al. (2002), explicitly modeled the cessation lag concept, using an analysis stratified on time since last exposure. All the other studies included in their models some estimate of latency, i.e., the delay between the critical exposure and diagnosis of disease or death. While not the same as the cessation lag, information about latency can also help inform our estimate for a cessation lag.

EXHIBIT 4
SUMMARY OF FINDINGS – LATENCY / CESSATION LAG ASSOCIATED WITH BENZENE-INDUCED LEUKEMIA

Study	Lags Tested	Findings
Silver et al. (2002)	Time since last exposed: 0; 0.01-4.9; 5-19.9; and >20 years	Generated SMRs for yearly follow-ups of Pliofilm Cohort starting in 1940 and extending from 1950 through 1996. Used Cox models to estimate effect of follow-up time on risk estimates. Stratified analysis of time since last exposed to benzene suggests that for this cohort, relative risk peaks in the first few years after cessation of exposure and that exposures 5-10 years prior to the cutoff have the most impact on risk. The results suggest that ensuring maximum protection for benzene workers requires assessing risk at its peak of 5-10 years since exposure.
Finkelstein (2000)	Exposure windows: 1-4, 5-9, 10-14, 15-19, 20-24, and 25-29 years before death	Case-control study in which the exposures of subjects with leukemia and matched controls were compared at various times before the death of the case. Looked backward from the date of death of the case subjects and compared the exposures of case and control subjects in specific exposure windows prior to the death of cases. Found no significant difference in the benzene exposures of subjects with leukemia and their matched controls 15 or more years prior to death of case. The highest risk was related to exposures incurred in the previous 10 years.
Hayes et al. (1997)	Recent (1.5-10 years) Distant (10+years)	Study of Chinese Worker Cohort that partitioned cumulative exposure into recent (1.5-10 years earlier) and distant (10 or more years earlier) exposure. Risk of ANLL/MDS was positively associated with recent benzene exposure, and additional distant exposure did not appear to further increase risk.
Crump (1994)	0, 3, 5 years	For the best-fitting class of risk models (multiplicative risk models using cumulative exposure), 5-year lag provided best fit to the data. Multiplicative risk models using weighted exposure generated estimates of latency for leukemia deaths of 6.7 yrs (AMML) and 7.7 years (all leukemia).
Rinsky et al. (2002)	0, 2.5, 5, and 10 years	Follow-up analysis of the Pliofilm Cohort (extended follow-up an additional 15 years). Study included at least 20 years of follow-up for every member. Model fit worsened with increasing lag. Zero lag linear model showed best fit, though 2.5 year lag only slightly less suitable. No data shown for longer lags.
Glass et al., 2004	≤15, >15 years	Nested case-control study of Australian petroleum workers (Health Watch cohort). Found that leukemia was most strongly associated with benzene exposures within 15 years of diagnosis; exposures more than 15 years prior to diagnosis showed little impact on risk.
Rushton and Romaniuk (1997)	0, 5, 10 years	Case-control study of petroleum workers in the UK. For all leukemia, risks did not change substantially with increasing lag. For AML, odds ratios for categories of cumulative exposure tended to increase with

		increasing lag, model fit tended to improve.
Schnatter et al. (1996a)	0, 5 years	Case-control study of petroleum workers in the Canada. Effect of lag on risk estimates was inconsistent.
Guenel et al. (2002)	2, 5, 10 years	Case-control study of utility workers in France. Results largely similar for different lags.
Bloemen et al. (2005)	0, 15 years	Cohort study of chemical workers assessing leukemia mortality rate. Lagging exposure by 15 years did not increase risk estimates.

Estimates of latency vary across studies. In general, most studies in Exhibit 4 found that latency estimates of 10 years or fewer fit the data best. Studies of the Pliofilm Cohort (Crump, 1994, Rinsky et al., 2002) tended to find slightly lower latency estimates, while Hayes et al. (1997) study of the Chinese Worker Cohort found stronger effects of "recent" exposures, where recent was defined as between 1.5 and 10 years prior to diagnosis. Finkelstein (2000) used the Pliofilm cohort dataset to compare exposures of leukemia cases and controls in specific exposure windows prior to the death of the case. He also found that the highest risk was related to exposures within the last 10 years prior to death, and that there was no significant difference in exposures between cases and controls 15 or more years prior to death. The case-control analysis by Glass et al. (2004) also found that exposures more than 15 years prior to diagnosis had little impact on leukemia risk. No other study found evidence suggesting a latency period longer than 15 years.

Silver et al. (2002) re-analyzed the Rinsky et al. (2002) Pliofilm Cohort dataset, generating SMRs for yearly follow-ups from 1950 through 1996. Silver et al. then analyzed these data stratified on time since last exposure and found that leukemia risk peaks within the first five years following cessation of exposure. He also found, in a separate analysis of exposure windows, that exposures five to ten years prior to the cutoff have the maximum impact on risk, and that exposures between ten and 15 years prior to cutoff may also contribute to a lesser degree. However, the authors of this study note that the smaller number of cases from the Pliofilm Cohort limits the precision with which they can define the relative risks in each period.

IMPLICATIONS FOR ANALYTICAL PLAN

This section discusses the implications of the findings of our literature review for the analytical plan for the benzene case study. We divide our conclusions and recommendations into those affecting cancer endpoints and those affecting non-cancer endpoints.

Cancer Endpoints

Based on the results of our literature review on the health effects of benzene exposure, and evidence gathered by EPA in the IRIS support document for benzene carcinogenicity, we propose to quantify the avoided cases of leukemia due to changes in benzene exposure through a dose-response analysis. We prefer to use the outcome of total leukemia for the primary estimate, since this endpoint is the most data rich, compared to the limited evidence for a link with benzene and the specific leukemia types (AML, ALL, CML and CLL). However, EPA may wish to consider conducting a sensitivity analysis that estimates avoided cases of AML, since this subtype has most evidentiary support among the different types of leukemia.

The two strongest cohort studies examining the link between benzene and leukemia have different strengths and limitations. However, the IRIS profile for benzene currently supports the use of data from the Pliofilm cohort for calculating potency estimates. Therefore, we propose to use beta coefficients reported by Crump (1994) for our primary estimate of avoided leukemias, as indicated in our analytical plan. We propose to use risk estimates based on the cumulative exposure linear multiplicative risk model presented in Crump (1994). We are not proposing to incorporate a threshold, because we do not find current evidence on potential thresholds for benzene-induced leukemia to be persuasive. In addition, although there is growing evidence supporting a supralinear dose-response function, there does not appear to be enough conclusive evidence to depart from the default linear low-dose extrapolation as discussed in EPA's *Guidelines for Carcinogen Risk Assessment* (EPA, 2005).

Despite its limitations, the Chinese Worker Cohort data has certain advantages over the Pliofilm Cohort, such as large sample size and benzene exposure levels that are more consistent with ambient exposures. Therefore, we could perform a sensitivity analysis using the results of the Chinese Worker Cohort. The California Environmental Protection Agency (CalEPA) recently used the Chinese Worker Cohort data in calculating a Public Health Goal for benzene (CalEPA, 2001). The CalEPA analysis of dose-response in the Chinese Worker Study could serve as the basis for our sensitivity analysis. In their analysis, the authors assumed a linear dose-response function for extrapolation to low doses. We agree with this conclusion because EPA's *Guidelines for Carcinogen Risk Assessment* (EPA, 2005) state that linear extrapolation should be used when the mode of action is uncertain, which is the case for benzene. In addition, given the low concentrations that are likely to be experienced in our case study, a linear approximation may be a reasonable fit, even if the overall dose-response function in supralinear, provided the data from which the extrapolation is being made are not in the plateau region of the curve.[8] Due to the growing body of evidence for supralinearity, even potentially at low doses (Rappaport et al., 2002), we could consult with the Office of Research and Development (ORD) on the usefulness of and level of effort needed to develop an alternate supralinear model for the Chinese Worker Cohort data as part of the sensitivity analysis.

In our previous analytical plan, we proposed assuming a 5-year lag between benzene exposure and leukemia as a first estimate of the cessation lag that determines the temporal distribution of benefits. Our literature search has discovered evidence that longer lag periods might also be valid, though the majority of the literature suggests that most cases would occur within 10 years, with some smaller number of cases occurring between 10 and 15 years. The Silver et al. (2002) study in particular specifically addresses the cessation lag concept and finds results suggesting that while mean latency may be in the five to ten year range, the move towards a new steady state of risk may begin fairly quickly, and a significant portion of deaths due to past exposures may occur within the first five years following a change in exposure. This finding, combined with the lag results from other studies points towards a lag structure where a new steady-state risk level is reached within 15 years following a regulatory change. Within this 15-year period, most of the risk reduction will be realized between five and ten years post-change, with smaller risk reductions accruing within the first five years and within 10 to 15 years

---

[8] In this case, the linear slope might be too shallow, underestimating the true dose-response relationship at low doses. To address this, the CalEPA analysis excluded data points expected to be in the plateau region of the curve.

following the change. Identifying reasonable assumptions for distributing the risk reductions across and within these periods will require additional study of the Silver et al. paper and consultation with its authors.

In addition to leukemia, benzene exposure has been associated with other cancerous health endpoints in epidemiologic studies, such as HL and NHL (Hayes et al., 1997), multiple myeloma (Rinsky et al., 1987 & 2002; Wong et al., 1995), and MDS (Hayes et al., 1997) but data on these endpoints are inconsistent and do not yet support a quantitative evaluation. We propose to describe the evidence for associations of benzene with these endpoints qualitatively.

Non-Cancer Endpoints

The dose-response data underlying the RfC (Rothman, 1996a) do not support a fully quantitative estimate of avoided "cases" of reduced lymphocytes expected at environmental levels due to the small number of data points (two). However, recent studies by Lan et al. (2004) and Qu et al. (2002), may support this effort, since they provide three and four data points, respectively, from which it may be possible to extrapolate a dose-response relationship. Other strengths of these studies include large number of exposed cases (250 and 130), detailed exposure assessment (measured), control for confounding factors, and exposure measurements below 1 ppm, which would allow for better low-dose extrapolation. Thus, quantification of "cases" may be possible, though we recommend consulting with ORD to discuss the level of effort required to pursue this approach.

Another factor to consider in deciding whether to quantify cases is the uncertain health impact of reduced lymphocytes, which would likely make it difficult to monetize such effects. The IRIS profile states that decreased lymphocyte count is a biomarker of exposure and is also thought to have a potential role as a "sentinel" effect (i.e., an early sign of toxicity in the bone marrow), but the effect itself is of uncertain clinical significance to the average population. The significance of the effect depends on both the magnitude of the decrease in lymphocytes and an individual's baseline lymphocyte level. For example, the effect of reduced lymphocytes might be more significant for individuals whose immune systems were compromised (e.g., those suffering from HIV/AIDS). Because of uncertainty in the impact on average healthy individuals, we expect we may be unable to value these avoided "cases" of reduced lymphocytes.

An alternative approach, outlined in our original analytical plan, could be to assess this endpoint by reporting the difference in the number of individuals experiencing benzene concentrations above the RfC under the pre-CAAA and post-CAAA scenarios. While we recognize that exposure above the RfC does not necessarily imply the presence of an adverse effect in a given individual, this estimate nonetheless provides some measure of progress towards reducing the likelihood of adverse hematological effects.

REFERENCES

Adegoke, O.J., Blair, A., et al. (2003). "Occupational history and exposure and the risk of adult leukemia in Shanghai." Annals of Epidemiology 13:485-494.

Bergsagel, D.E., Wong, O., et al. (1999). "Benzene and multiple myeloma: Appraisal of the scientific evidence." Blood 94(4):1174-1182.

Bezabeh, S., Engel, A., et al. (1996). "Does benzene cause multiple myeloma? An analysis of the published case-control literature." Environmental Health Perspectives 104(Suppl 6):1393-1398.

Bloemen, L.J., Youk, A., et al. (2004). "Lymphohaematopoietic cancer risk among chemical workers exposed to benzene." Occupational and Environmental Medicine 61:270-274.

Bogadi-Sare, A., Zavalic, M., et al. (2000). "Study of some immunological parameters in workers occupationally exposed to benzene." International Archives of Occupational and Environmental Health 73:397-400.

Budinsky, R.P., DeMott, R.P., et al. (1999). "An evaluation of modeled benzene exposure and dose estimates published in the Chinese-National Cancer Institute Collaborative epidemiology studies." Regulatory Toxicology and Pharmacology 30:244-258.

California Environmental Protection Agency (2001). "Public Health Goal for benzene in drinking water." Office of Environmental Health Hazard Assessment.

Collins, J.J., Conner, P., et al. (1991). "A study of the hematologic effects of chronic low-level exposure to benzene." Journal of Occupational Medicine 33:619-626.

Collins, J.J., Ireland, B.K., et al. (1997). "Evaluation of lymphopenia among workers with low-level benzene exposure and the utility of routine data collection." Journal of Occupational and Environmental Medicine 39(3):232-237.

Collins, J.J., Ireland, B.K., et al. (2003). "Lymphohaematopoeitic cancer mortality among workers with benzene exposure." Occupational and Environmental Medicine 60:676-679.

Costantini, A.S., Quinn, M., et al. (2003). "Exposure to benzene and risk of leukemia among shoe factory workers." Scandanavian Journal of Work and Environmental Health 29(1):51-59.

Crump, K.S. (1994). "Risk of benzene-induced leukemia: A sensitivity analysis of the Pliofilm Cohort with additional follow-up and new exposure estimates." Journal of Toxicology and Environmental Health 42:219-242.

Crump, K.S. (1996) "Risk of benzene-induced leukemia predicted from the Pliofilm Cohort." Environmental Health Perspectives 104(Suppl 6): 1437-1441.

Crump, K.S., and Allen, B.C. (1984) "Quantitative estimates of risk of leukemia from occupational exposure to benzene." US Department of Labor, Washington DC (OSHA Docket H-059b, Exhibit 152, Annex B).

DeCoufle, P., Blatter, W.A., et al. (1983). "Mortality among chemical workers exposed to benzene and other agents." Environmental Research 30:16-25.

Dosemeci, M., Li, G.-L., et al. (1994). "Cohort study among workers exposed to benzene in China: II. Exposure assessment." American Journal of Industrial Medicine 26:401-411.

Dosemeci, M., Yin, S.N., et al. (1996). "Indirect validation of benzene exposure assessment by association with benzene poisoning." Environmental Health Perspectives 104(Suppl 6): 1343-1347.

Finkelstein, M.M. (2000). "Leukemia after exposure to benzene: Temporal trends and implications for standards." American Journal of Industrial Medicine 38:1-7.

Gerstman, B.B. (1998) Epidemiology Kept Simple: An Introduction to Classic and Modern Epidemiology. John Wiley & Sons Publishing, New York, NY.

Glass, D.C., Gray, C.N., et al. (2003). "Leukemia risk associated with low-level benzene exposure." Epidemiology 14(5):569-577.

Glass, D.C., Sim, M.R., et al. (2004). "Leukemia risk and relevant Benzene exposure period – Re: Follow-up time on risk estimates, Am J Ind Med 42:481-489, 2002." Am J Ind Med 45 :222-223.

Guenel, P., Imbernon, E., et al. (2002). "Leukemia in relation to occupational exposures to benzene and other agents: A case-control study nested in a cohort of gas and electric utility workers." American Journal of Industrial Medicine 42:87-97.

Hayes, R.B., Yin, S.-N., et al. (1997). "Benzene and the dose-related incidence of hematologic neoplasms in China." Journal of the National Cancer Institute 89(14):1065-1071.

Hayes, R.B., Yin, S-N., et al. (2000). "Benzene and lymphohematopoietic malignancies in China." Journal of Toxicology and Environmental Health 61:419-432.

Hayes, R.B., Songnian, Y., et al. (2001). "Benzene and lymphohematopoietic malignancies in humans." American Journal of Industrial Medicine 40:117-126.

Ireland, B., Collins, J.J., et al. (1997). "Cancer mortality among workers with benzene exposure." Epidemiology 8:318-320.

Khuder, S.A., Youngdale, M.C., et al. (1999). "Assessment of complete blood count variations among workers exposed to low levels of benzene." Journal of Occupational and Environmental Medicine 41(9):821-826.

Lamm, S.H., Engel, A., et al. (2004) "Non-Hodgkin's lymphoma and benzene exposure - A systematic literature review." Abstract from the Recent Advances in Benzene Toxicity conference in Munich, Germany in October, 2004.

Lamm, S.H., Engel, A., et al. (2005) "Non-Hodgkin's lymphoma and benzene exposure - A systematic literature review." Chemico-biological Interactions 153-154:231-237.

Lan, Q, Zhang, L., et al. (2004). "Hematotoxicity in workers exposed to low levels of benzene." Science 306: 1774-1776.

Paustenbach, D.J., Price, P.S., et al. (1992). "Reevaluation of benzene exposure for the Pliofilm (rubberworker) Cohort (1936-1976)." Journal of Toxicology and environmental Health 36:177-231.

Paxton, M.B., Chinchili, V.M., et al. (1994). "Leukemia risk associated with benzene exposure in the Pliofilm Cohort. II. Risk estimates." Risk Analysis 14(2):155-161.

Qu, Q., Shore, R., et al. (2002). "Hematological changes among Chinese workers with a broad range of benzene exposures." American Journal of Industrial Medicine 42:275-285.

Rappaport, S.M., Waidyanatha, S., et al. (2002). "Albumin adducts of benzene oxide and 1,4-benzoquinone as measures of human benzene metabolism." Cancer Research 62:1330-1337.

Rinsky, R.A., Hornung, R.W., et al. (2002). "Benzene exposure and hematopoietic mortality: A long-term epidemiologic risk assessment." American Journal of Industrial Medicine 42:474-480.

Rinsky, R.A., Smith, A.B., et al. (1987). "Benzene and leukemia: An epidemiologic risk assessment." New England Journal of Medicine 316:1044-1050.

Rinsky, R.A., Young, R.J., et al. (1981). "Leukemia in benzene workers." American Journal of Industrial Medicine 2(3):217-45.

Rothman, N, Smith, M.T., et al. (1996b). "An epidemiologic study of early biologic effects of benzene in Chinese workers." Environmental Health Perspectives 104(Suppl 6):1365-1370.

Rothman, N., Haas, R., et al. (1995). "Benzene induces gene-duplicating but not gene-inactivating mutations at the glycophorin A locus in exposed humans." Proceedings of the National Academy of Sciences 92:4069-4073.

Rothman, N., Li, G.-L., et al. (1996a). "Hematotoxicity among Chinese workers heavily exposed to benzene." American Journal of Industrial Medicine 29:236-246.

Rothman, N., Smith, M.T., et al. (1997). "Benzene poisoning, a risk factor for hematological malignancy, is associated with the NQ01^{609}C-T mutation and rapid fractional excretion of chlorzoxazone." Cancer Research 57:2839-2842.

Rothman, N., Bechtold, W.E., et al. (1998). "Urinary excretion of phenol, catechol, hydroquinone, and muconic acid by workers occupationally exposed to benzene." Occupational and Environmental Medicine 55:705-711.

Rusthton, L. and Romaniuk, H. (1997). "A case-control study to investigate the risk of leukaemia associated with exposure to benzene in petroleum marketing and distribution workers in the United Kingdom." Occupational and Environmental Medicine 54(3):152-166.

Savitz, D.A., and Andrews, K.W. (1997). "Review of epidemiologic evidence on benzene and lymphatic and hematopoietic cancers." American Journal of Industrial Medicine 31:287-295.

Schnatter, R.A., Armstrong, T.W., et al. (1996a). "Lymphohaematopoietic malignancies and quantitative estimates of exposure to benzene in Canadian petroleum distribution workers." Occupational and Environmental Medicine 53(11):773-781.

Schnatter, R.A., Nicolich, M.J., et al. (1996b). "Determination of leukemogenic benzene exposure concentrations: Refined analyses of the Pliofilm Cohort." Risk Analysis 16(6):833-840.

Silver, S.R., Rinsky, R.A., et al. (2002). "Effect of follow-up time on risk estimates: A longitudinal examination of the relative risks of leukemia and multiple myeloma in a rubber hydrochloride cohort." American Journal of Industrial Medicine 42:481-489.

Sonoda, T., Nagata, Y., et al. (2001). "Meta-analysis of multiple myeloma and benzene exposure." Journal of Epidemiology 11(6):249-54.

Sorahan, T., Kinlen, L.J., et al. (2005). "Cancer risks in a historical UK cohort of benzene exposed workers." Occupational and Environmental Medicine 62:231-236.

Swaen, G.M.H, Scheffers, T., et al. (2005). "Leukemia risk in Caprolactam workers exposed to benzene." Annals of Epidemiology 15:21-28.

U.S. EPA (1998). "Carcinogenic effects of benzene: An update." Office of Research and Development, Washington, DC. EPA/600/P-97/001F.

U.S. EPA (2002). "Toxicological review of benzene (non-cancer effects)." Office of Research and Development, Washington DC. EPA/635/R-02/001F.

U.S. EPA. (2004). " Advisory on Plans for Health Effects Analysis in the Analytical Plan for EPA's Second Prospective Analysis - Benefits and Costs of the Clean Air Act, 1990-2020, Advisory by the Health Effects Subcommittee of the Advisory Council for Clean Air Compliance Analysis." Science Advisory Board, Washington, D.C. EPA-SAB-COUNCIL-ADV-04-002.

U.S. EPA. (2005). "Guidelines for carcinogen risk assessment." Risk Assessment Forum, Washington, D.C. EPA/630/P-03/001F.

Utterback, D.F., Rinsky, R.A. (1995). "Benzene exposure assessment in rubber hydrochloride workers: a critical evaluation of previous estimates." American Journal of Industrial Medicine 27(5):661-76.

Ward, E., Hornung, R., et al. (1996). "Risk of low red or white blood cell count related to estimated benzene exposure in a rubberworker cohort (1940-1975)." American Journal of Industrial Medicine 29:247-257.

Wong, O. (1995). "Risk of acute myeloid leukaemia and multiple myeloma in workers exposed to benzene." Occupational and Environmental Medicine 52:380-384.

Wong, O. (1999). "A critique of the exposure assessment in the epidemiologic study of benzene-exposed workers in China conducted by the Chinese Academy of Preventative Medicine and the US National Cancer Institute." Regulatory Toxicology and Pharmacology 30:259-267.

Wong, O., Raabe, G.K. (2000). "Non-Hodgkin's lymphoma and exposure to benzene in a multinational cohort of more than 308,000 petroleum workers, 1937 to 1996." Journal of Occupational and Environmental Medicine 42(5):554-568

Yin, S.-N., Hayes, R.B., et al. (1996). "An expanded cohort study of cancer among benzene-exposed workers in China." Environmental Health Perspectives 104(Suppl 6):1339-1341.

Yin, S.-N., Li, G.-L., et al. (1987). "Leukemia in benzene workers: A retrospective cohort study." British Journal of Industrial Medicine 44:124-128.

ATTACHMENT 1 – ADDENDUM TO APPENDIX C

In order to judge the weight-of-evidence available for benzene-related health effects and the appropriateness for including these effects in the case study, we relied upon the criteria listed below. We reviewed the body of evidence linking each health effect considered. The weight-of evidence rating was IEc's subjective assessment based on the extent to which the link between a particular health effect and benzene exposure met the criteria in Table 1 below. Exhibit 2 of this appendix provides a summary of our assessment of the weight-of-evidence for each health effect investigated.

Table 1. Criteria for Judging the Weight-of-Evidence for Benzene-Related Health Effects

- Human epidemiological studies (e.g., occupational) exist examining the link between the health effect and benzene exposure.
- The body of evidence includes studies with a cohort design, which is preferable to other types of study designs, such as case-control.
- The studies are published in peer reviewed journals or have otherwise undergone an external peer review process.
- Consistent positive associations are found between the health effect and benzene exposure. Preferably, these associations are statistically significant.
- The studies have sufficient study population and number of cases to produce stable estimates (with relatively narrow confidence bounds).
- The studies are well designed, with few methodological issues (e.g., limited confounding by exposure to other chemicals).
- The studies include measured benzene exposures, preferably through individual monitoring.
- The benzene exposures experienced by the subjects in the studies are similar to ambient benzene levels.
- The studies have been re-analyzed by independent parties and the results verified.

APPENDIX D:

LIFE-TABLE MODEL EQUATIONS

APPENDIX D | LIFE TABLE MODEL EQUATIONS

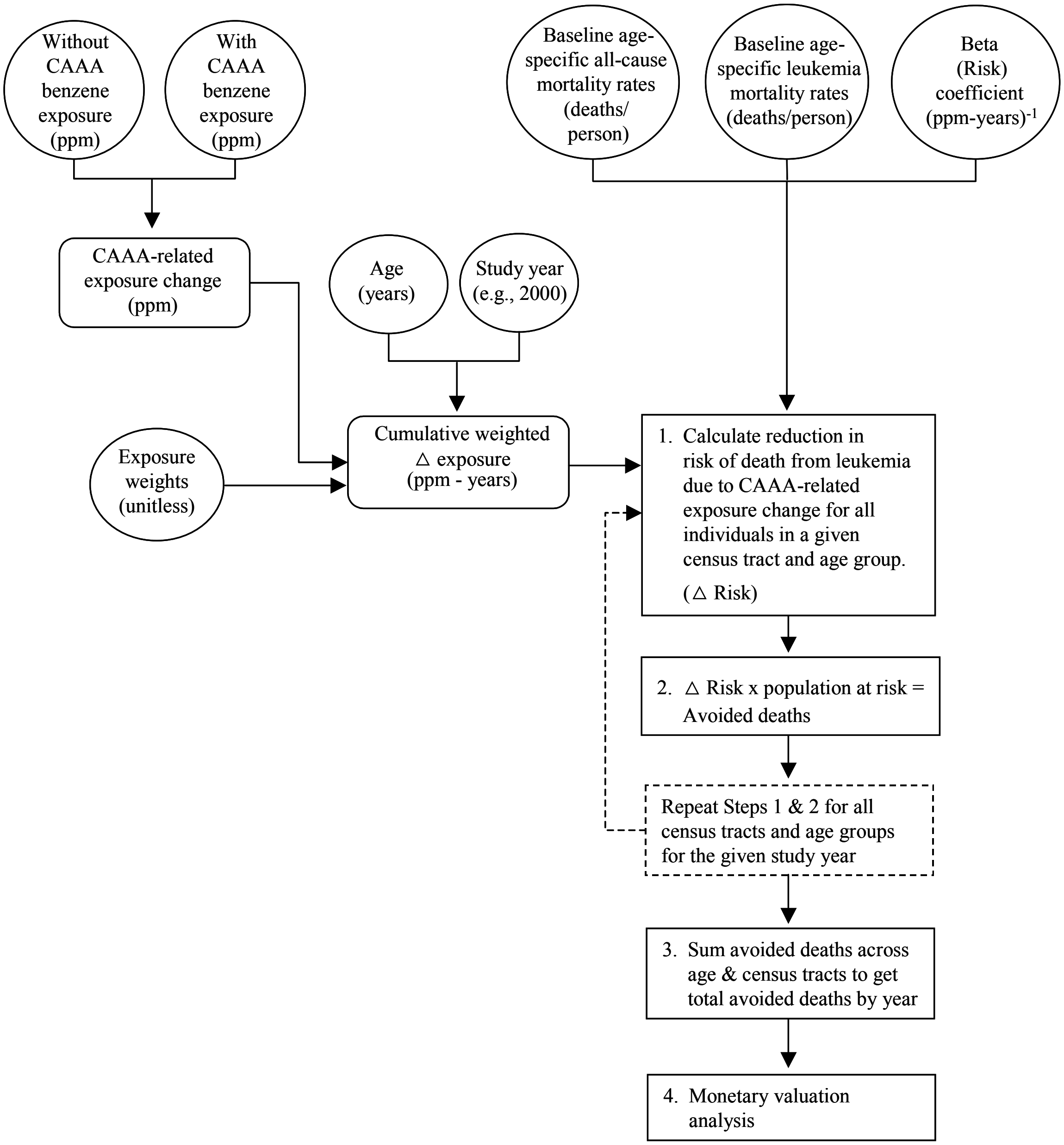
Figure D-1 provides an overview of the life table model.[1] The model involves calculating cumulative exposure estimates for each five-year age group in each census tract in each study year, which consists of a sum of previous exposure. The previous exposures are weighted differentially, depending on their influence on leukemia mortality rates. These cumulative weighted exposures are used to calculate the difference in risk of dying from leukemia between the *With-* and *Without-Clean Air Act Amendment* (CAAA) scenarios. The risk calculations are then repeated for each census tract and age group combination. The resulting risk values are then multiplied by the population in that census tract to calculate an estimate of avoided deaths from leukemia. Next, we summed avoided deaths across all age groups and census tracts to calculate an estimate of total cumulative avoided deaths by study year and across the entire study period. We then used the estimates of avoided deaths to calculate the monetary benefits related to CAAA-related reductions in benzene exposure.

The model begins with the raw exposure data from HAPEM6 and creates a cumulative weighted exposure measure for each age group in each census tract for each study year (e.g., 2000). This was done by first subtracting each raw five-year average exposure value under the *With-CAAA* scenario for each five-year age group in each census tract from the raw five-year average exposure value under the *Without-CAAA* scenario for the same five-year age group in the same census tract to get a “delta exposure” value, which represents the CAAA-related exposure change.

---

[1] Figure D-1 and equations presented below assume that the model is run with leukemia mortality rates. The model can also be run with leukemia incidence rates, using the same dose-response slope factor. The difference between these two runs represent an estimate of non-fatal cases of leukemia.

FIGURE D-1: LIFE-TABLE MODEL OVERVIEW



Note: This flowchart assumes the model is being run with leukemia mortality data. The model can also be run with leukemia incidence data. The difference between the model results for these two runs represents an estimate of avoided non-fatal cases of leukemia.

Equation D-1: $\Delta e_{i,j,k} = n_{i,j,k} - c_{i,j,k}$

Where:

$\Delta e_{i,j,k}$ (ppm) = the difference in raw exposure between the *Without-CAAA* scenario and the *With-CAAA* scenario for age group i in period j in census tract k;

$n_{i,j,k}$ (ppm) = raw exposure value under the *Without-CAAA* scenario for age group i in period j in census tract k; and

$c_{i,j,k}$ (ppm) = raw exposure value under the *With-CAAA* scenario for age group i in period j in census tract k.

We then created a historical exposure profile for each age group for each five-year period in each census tract to get a cumulative weighted exposure value representing the difference between the *With*- and *Without-CAAA* scenarios for each age in each five-year period in each census tract.

Equation D-2: $\Delta E_{i,j,k} = \Sigma\ (\Delta e_{i,j,k} \times w_t)$

$j = \max(j-i^{*},1995)$ to j; $i = 0$ to i; and $t = 0$ to $j - \max(j-i,1995)$

* i represents the starting age of the age group. For example, age group 5 includes those aged 5-9.

Where:

$\Delta E_{i,j,k}$ (ppm-years) = cumulative weighted exposure representing the difference between the *With*- and *Without-CAAA* scenarios for age group i in period j in census tract k;

$\Delta e_{i,j,k}$ (ppm) = raw exposure data representing the difference between the *With*- and *Without-CAAA* scenarios for age group i in period j in census tract k; and

w_t (unitless) = weight corresponding to a given value of t.[2]

---

[2] The weighting function took on the following form: $w(t) = (t/K^2) \exp(-t/K)$. Where: t = the number of years prior to the current year; and K = number of years prior to the current year when the weight reaches its maximum (this also represents the latency estimate).

We then combined the cumulative weighted Δ exposure calculated above with baseline all-cause and leukemia mortality rates and the dose-response slope factor from the selected epidemiologic study to calculate the risk of dying from leukemia in a given five-year period. Equations D-3 and D-4 below are a function of the relative ratio of leukemia deaths to all deaths and the probability of dying in a given five-year period, conditional on survival up to the five-year period for a Baseline scenario (no additional benzene exposure) or an Exposed scenario (with additional exposure to benzene).

Baseline

Equation D-3: $R^o_i = \alpha_i/\delta_i \times S(1,i) \times (1- q_i)$

Where:

R^o_i = baseline risk of leukemia in the absence of additional benzene exposures for age group i;

α_i (deaths/person) = baseline leukemia mortality rate for age group i (county-specific);

δ_i (deaths/person) = baseline all-cause mortality rate for age group i (county-specific);

q_i = probability of surviving through age group i = $(\exp(-5 \times \delta_i))$;

$1-q_i$ = probability of dying while in age group i;

$S(1,i)_j$ = probability of surviving up to age group i in period j. This is the product of the probabilities of surviving each prior age $(q_1 \times q_2 \times \ldots \times q_{i-1} = S(1,i))$ with $S(1,1) = 1.0$. Can be calculated by multiplying $S_{i-1,j-1}$ and q_{i-1}.

Exposed

Equation D-4: $R^e_{i,j,k} = h_{i,j,k}/h_{i,j,k}{*} \times S(1,i)_{j,k} \times (1- q_{i,j,k})$

Where:

$R^e_{i,j,k}$ = risk of leukemia due to benzene exposure for age group i in period j in census tract k;

$h_{i,j,k}$ (deaths/person) = exposed leukemia mortality rate for age group i in period j in census tract k = $\alpha_i\ (1 + \beta\ \Delta E_{i,j,k})$;

Where:

α_i (deaths/person) = baseline leukemia mortality rate for age group i (county-specific);

β (ppm-years)$^{-1}$ = risk coefficient from epidemiologic study;

$\Delta E_{i,j,k}$ (ppm-years) = difference between the cumulative weighted exposure for the *With*- and *Without-CAAA* scenarios for age group i in period j in census tract k;

$h_{i,j,k}$* (deaths/person) = exposed all-cause mortality rate for age group i in period j in census tract k = $\delta_i + (h_{i,j,k} - \alpha_i)$;

Where:

δ_i (deaths/person) = baseline all-cause mortality rate for age group i (county-specific).

$q_{i,j,k}$ = probability of surviving through age group i in period j in census tract k = $(\exp(-5 \times h_{i,j,k}^*))$;

1-$q_{i,j,k}$ = probability of dying while in age group i in period j in census tract k ; and

$S(1,i)_{j,k}$ = probability of surviving up to age group i in period j in census tract k.. This is the product of surviving each prior age group ($q_{0,\max(j-i,1995),k} \times q_{1,\max(j-i,1995)+1,k} \times \ldots \times q_{i-1,j-1,k} = S(1,i)_{j,k}$) with $S(1,1)_{j,k} = 1.0$. Can be calculated by multiplying $S_{i-1,j-1,k}$ and $q_{i-1,j-1,k}$.

To calculate the risk due to the additional benzene exposures experienced under the *Without-CAAA* scenario, we subtracted the baseline risk from the exposed risk, using Equation D-5.

Equation D-5: $R^e_{i,j,k} - R^o_i = \Delta R_{i,j,k}$

Where:

$R^e_{i,j,k}$ = risk of leukemia due to benzene exposure for age group i in period j in census tract k;

R^o_i = baseline risk of leukemia in the absence of additional benzene exposures for age group i; and

$\Delta R_{i,j,k}$ = risk of dying from leukemia due to CAAA-related exposures for all individuals in age group I in period j in census tract k.

APPENDIX E:

ATTACHED GARAGE ANALYSIS EQUATIONS

APPENDIX E | ATTACHED GARAGE ANALYSIS EQUATIONS

Our approach for assessing the rough magnitude of additional potential benefits that may result from Clean Air Act Amendment (CAAA)-related reductions of in-garage benzene emissions in 2020 involved three steps: first, we assessed the percent reduction in total emissions occurring within attached garages due to the CAAA in 2020; second, we applied the percent reduction to an estimate of average benzene exposure attributable to attached garages; third, we calculated the annual number of avoided cases of leukemia in the Houston area in 2020 that would be expected based on the CAAA-related reduction in attached garage-related exposures. We describe these steps in greater detail below.

Step 1

We first calculated the percent reduction in total emissions occurring within attached garages due to the CAAA in 2020. Total emissions include emissions from both non-road and on-road source categories. We calculated the difference using the following equation:

$$R_g = (E_{g\ \mathit{Without\text{-}CAAA}} - E_{g\ \mathit{With\text{-}CAAA}})/E_{g\ \mathit{Without\text{-}CAAA}}$$

Where:

R_g = percent reduction in emissions occurring within attached garages due to the CAAA;

$E_{g\ \mathit{Without\text{-}CAAA}}$= total emissions occurring within attached garages under the *Without-CAAA* scenario in tons/year ($Eg_{\text{non-road}\ \mathit{Without}} + Eg_{\text{on-road}\ \mathit{Without}}$); and

$E_{g\ \mathit{With\text{-}CAAA}}$ = total emissions occurring within attached garages under the *With-CAAA* scenario in tons/year ($E_{g\ \text{non-road}\ \mathit{With}} + E_{g\ \text{on-road}\ \mathit{With}}$).

We employed different approaches for estimating the non-road and on-road component of emissions occurring within attached garages under each of the scenarios, because of differences in the available emissions data for these two source categories. We describe the two approaches in detail below.

Non-road Emissions Occurring Within Attached Garages

In order to calculate the non-road emissions occurring within attached garages under each of the two scenarios, we first identified only those non-road vehicles or equipment that we would expect to be kept in a garage. These included all residential lawn and gardening equipment as well as recreational non-road vehicles.[1] We then took estimates of benzene emissions in 2020 in tons per year for each of the selected non-road vehicles and equipment and split the emissions estimates into emission categories (i.e., exhaust, evaporative, refilling).[2] For example, we used the following equation to calculate the evaporative fraction of total emissions from a particular non-road source under the *Without-CAAA* scenario:

$$E_{\text{evap NR } Without} = f_{\text{evap NR } Without} \times E_{\text{NR } Without}$$

Where:

$E_{\text{evap NR } Without}$ = the non-road emissions that are evaporative under the *Without-CAAA* scenario in tons/year;

$f_{\text{evap NR } Without}$ = the fraction of non-road emissions that are evaporative under the *Without-CAAA* scenario; and

$E_{\text{NR } Without}$ = the total non-road emissions under the *Without-CAAA* scenario in tons/year.

We repeated this process for each combination of emissions category (evaporative, exhaust, refilling), and scenario (*With-CAAA*, *Without-CAAA*).

Next, we applied to each category a factor describing the fraction of those emissions expected to occur within an attached garage. For each category, we employed a range of percentages for each fraction, using values reported in Appendix 3A of the Regulatory Impact Analysis (RIA) for the Mobile Source Air Toxics Rule (MSAT) (USEPA, 2007; hereafter, the "MSAT RIA"). Table 2 provides the ranges of percentages we used for each of the emissions categories. We used theses values in the following equation to estimate the total non-road emissions expected to occur with in an attached garage:

$$Eg_{\text{NR } Without} = (E_{\text{evap NR } Without} \times f_{g \text{ evap}}) + (E_{\text{exh NR } Without} \times f_{g \text{ exh}}) + (E_{\text{refill NR } Without} \times f_{g \text{ refill}})$$

Where:

$E_{g \text{ NR } Without}$ = total non-road emissions occurring within attached garages under the *Without-CAAA* scenario in tons/year;

$E_{\text{evap NR } Without}$ = the non-road emissions that are evaporative under the *Without-CAAA* scenario in tons/year;

---

[1] If a particular type of non-road vehicles had more than one variety, we took an average across all varieties. For instance, we took an average of the emissions from 2-stroke rotary tillers and 4-stroke rotary tillers to estimate the in-garage emissions from an average rotary tiller.

[2] These data were provided by E.H Pechan and Associates (Pechan, 2008a).

$f_{g\ evap}$ = the fraction of evaporative emissions that occur within an attached garage;

$E_{exh\ NR\ Without}$ = the non-road emissions that are exhaust-related under the *Without-CAAA* scenario in tons/year;

$f_{g\ exh}$ = the fraction of exhaust-related emissions that occur within an attached garage;

$E_{refill\ NR\ Without}$ = the non-road emissions that are refilling-related under the *Without-CAAA* scenario in tons/year; and

$f_{g\ refill}$ = the fraction of refilling-related emissions that occur within an attached garage.

We repeated this process for the 2020 *With-CAAA* scenario. We then summed across all non-road vehicles and equipment to estimate the total emissions from this source category occurring within an attached garage under each scenario.

TABLE 2: ASSUMED FRACTIONS OF EMISSIONS FROM NON-ROAD GASOLINE EQUIPMENT AND VEHICLES OCCURRING WITHIN AN ATTACHED GARAGE

EMISSIONS CATEGORY	RANGE OF VALUES
Exhaust	0 – 2 percent
Evaporative	90 – 100 percent
Refilling-Related	25 – 75 percent

Source: Appendix 3A of the MSAT RIA, page 3-133, footnote u.

On-road Emissions Occurring Within Attached Garages

The available data for on-road emissions included the annual benzene emissions factors under the 2020 *With*- and *Without*-CAAA scenarios for emissions that are expected to occur within a garage. Pechan generated these factors using MOBILE6.2 (Pechan, 2008b). We focused on emissions related to light-duty gasoline vehicles (LDGV) and light-duty gasoline trucks with a loaded vehicle weight of 3,750 pounds and below (LDGT1). These emissions factors included idle (grams/min), start up (grams/start), hot soak (grams/trip end), diurnal (grams/day), resting loss (grams/day), and idle resting loss (grams/min). The emission factors were on a per vehicle basis. In order to estimate the total 2020 on-road emissions in tons per year that occur within an attached garage, we made a number of assumptions.

We first converted all of the emissions factors into units of tons/year/vehicle. This process differed depending on the specific emission factor. We assumed that on average, each vehicle would make two trips originating at the home and two trips ending at the home.[3] We also assumed that each vehicle would idle for one minute for every trip start and end at the home.

We then summed all of the emissions for each of the two vehicle types (LDGV and LDGT1) and took an average across them. We then estimated the total in-garage on-road emissions in the Houston area in 2020 by multiplying the average emissions in tons/year/vehicle by an estimate of the average number of vehicles per garage as well as an estimate of the number of attached garages in the Houston area.[4]

Step 2

Once we calculated the percent reduction in total emissions occurring within attached garages due to the CAAA in 2020, we applied it to an estimate of average indoor benzene exposure attributable to attached garages reported in Appendix 3A of the MSAT RIA to calculate an expected attached-garage related exposure reduction, using the following equation:[5]

$$ER_g = R_g \times E_g$$

Where:

ER_g = average attached garage-related indoor benzene exposure reduction due to the CAAA in $\mu g/m^3$;

R_g = percent reduction in emissions occurring within attached garages due to the CAAA in 2020;

E_g = average indoor benzene exposure estimate attributable to attached garages (1.2 $\mu g/m^3$; Table 3A-1, USEPA, 2007).[6, 7]

---

[3] This assumption was based on an estimate of average number of trips per person per day from the National Household Travel Survey (http://nhts.ornl.gov/).

[4] We estimated the average number of vehicles per household from an estimate of the total number of households in the US and the total number of vehicles in the US from the US Energy Information Administration website (http://www.eia.doe.gov/emeu/rtecs/nhts_survey/2001/). The number of attached garages in the Houston area was estimated by first dividing the total population of Harris, Galveston, and Brazoria counties by the average number of people per household in the Houston area (http://www.hellohouston.com/Census.Cfm) to calculate the total number of households. We then multiplied this by the fraction of households in the West South Central Census Region with attached garages from the Residential Energy Consumption Surveys (http://www.eia.doe.gov/emeu/recs/recs2001/detail_tables.html).

[5] The estimate of average indoor benzene exposure attributable to attached garages from the MSAT RIA incorporates an estimate of the fraction of the national population living in homes with attached garages (34.7 percent) from the Residential Energy Consumption Survey (RECS). We found that the estimate for the West South Central Census Region (which includes Texas) was similar. Therefore, we did not make any adjustments to the exposure estimate.

[6] We selected the estimate from Table 3A-1 of the MSAT RIA that was based on all studies except those conducted in Alaska due to a number of differences expected in the attached garage-related exposures between Alaska and Houston. For instance, the fuel in Alaska has atypically high benzene levels, the housing characteristics differ between these two locations, there could potentially be different types of vehicles and equipment found within garages in these locations, and cold starts likely contribute to benzene exposures in Alaska, whereas this would not be a factor in the Houston area.

This approach makes the conservative assumption that the percent reduction in in-garage emissions of benzene will result in an equivalent percent reduction in the component of indoor benzene exposure contributed by the attached garage.

Step 3

In the final step, we calculated the annual number of avoided cases of leukemia in the Houston area in 2020 that would be expected based on the CAAA-related reduction in attached garage-related benzene exposures, using the following equation:

$$\text{Annual Avoided Cases in 2020} = (ER_g \times IUR \times P)/LT$$

Where:

ER_g = average garage-related exposure reduction due to the CAAA in $\mu g/m^3$;

IUR = range of Inhalation Unit Risks for benzene in $(\mu g/m^3)^{-1}$;

P = total population in the Houston case study area; and

LT = lifetime, 70 years.

In this step, we multiply the exposure to the entire population in the Houston area because this exposure estimate represents a weighted average value across the population (see footnote 6).

REFERENCES

E.H. Pechan and Associates (2008a). *Nonroad Evaporative Fraction for IEcs.xls* [electronic file]. Transmitted to Henry Roman from Kirstin Thesing on January 25, 2008.

E.H. Pechan and Associates (2008b). *Garage_Benzene_Reducs_2020.xls* [electronic file]. Transmitted to Henry Roman and Tyra Walsh from Maureen Mullen on January 24, 2008.

U.S. EPA. (2007). Control of Hazardous Air Pollutants from Mobile Sources: Regulatory Impact Analysis. Office of Transportation and Air Quality. EPA420-R-07-002.

---

[7] This value represents a weighted average exposure across the population. It was calculated by multiplying the average indoor benzene concentration attributable to an attached garage by the fraction of the population living in a home with an attached garage and the time spent in a home with an attached garage.