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

Draft Report | March 14, 2008

INDUSTRIAL ECONOMICS, INCORPORATED

IEc

TABLE OF CONTENTS

LIST OF TABLES

IEc

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

IEc

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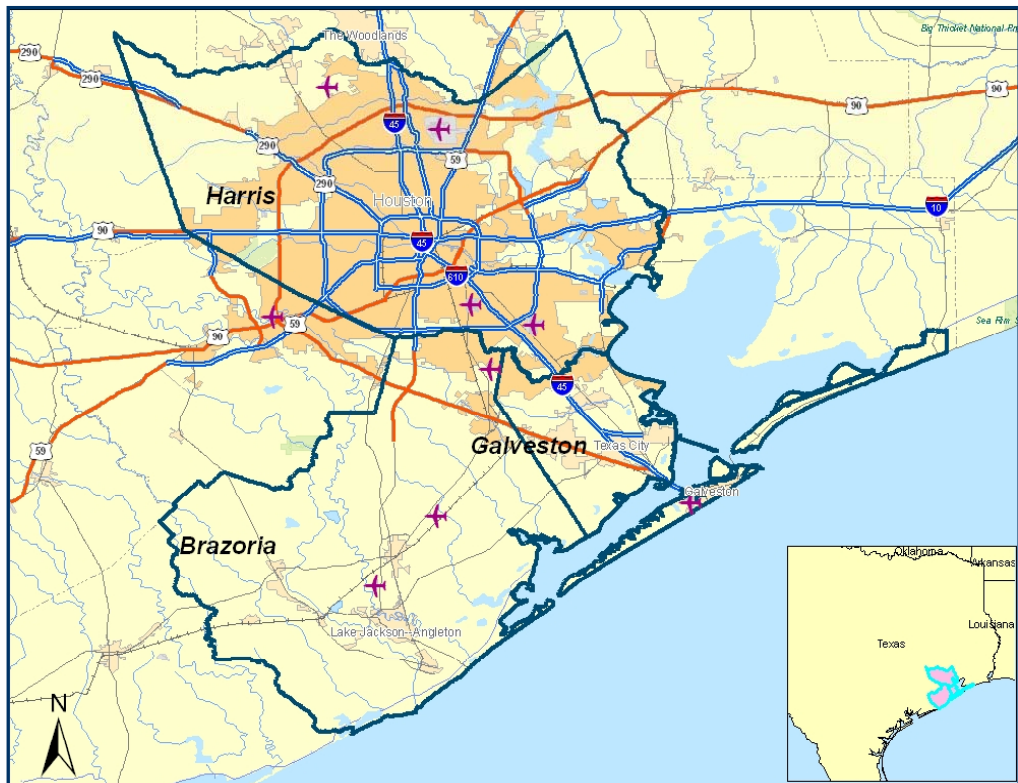
The timeframe for this analysis, 1990 through 2020, matches that used in the criteria pollutant analysis of the second prospective Section 812 study. In addition to the base year, 1990, we model results for three target years, 2000, 2010, and 2020.

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

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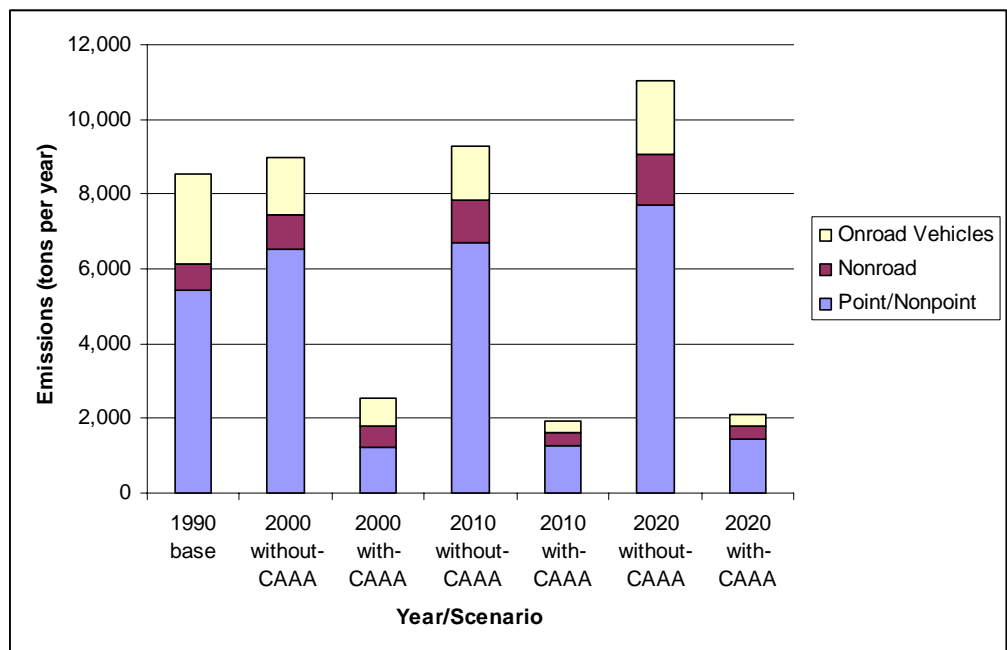
[1] Our modeling does not include indoor sources of exposure.

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.

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.
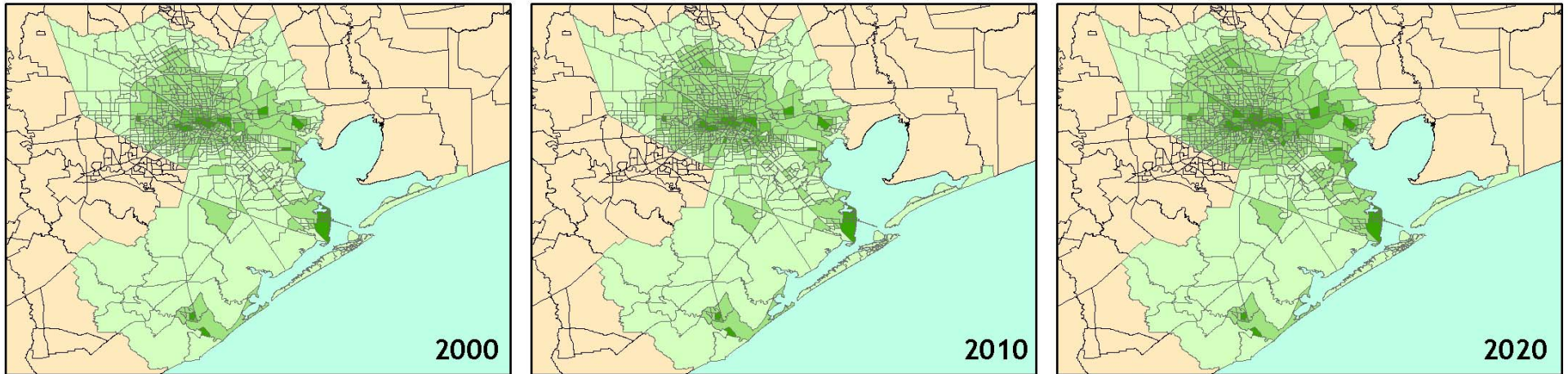
We focused our health benefits analysis on quantifying avoided cases of leukemia (all types), based on an extensive review of the available health effects literature. To estimate the avoided cases associated with benzene reductions in the study area, we constructed a life-table based risk assessment 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 the base value of statistical life (VSL) estimates used for particulate matter (PM)-mortality valuation, 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
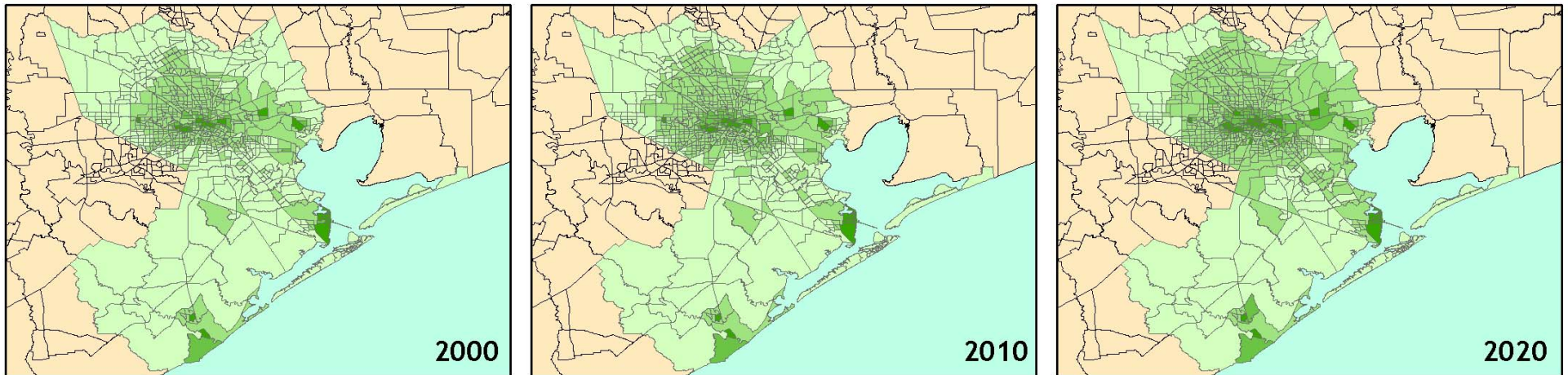
HAPEM RESULTS

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 cumulative total number of expected cases avoided from 1990 through each study year. 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: CUMULATIVE AVOIDED LEUKEMIA CASES (FATAL AND NON-FATAL) BY STUDY YEAR DUE TO CAAA-RELATED BENZENE EXPOSURE CHANGES IN THE HOUSTON AREA

STUDY YEAR	CUMULATIVE AVOIDED CASES OF LEUKEMIA		
	AVOIDED FATAL CASES	AVOIDED NON-FATAL CASES	TOTAL AVOIDED CASES
1990	0	0	0
2000	0.5	0.4	0.9
2010	2	2	4
2020	5	4	9

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

DISCOUNT PERCENTAGE	TOTAL BENEFITS (1990 NPV, MILLIONS OF1999$)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
Primary Estimate (5%)	$11	$2	$13
No Discounting	$29	$5	$34
Low Discount Rate (3%)	$16	$3	$19
High Discount Rate (7%)	$8.2	$1	$9

IEc

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

Total Cumulative Avoided Cases by Study Year - Primary Estimate and Sensitivity Analyses
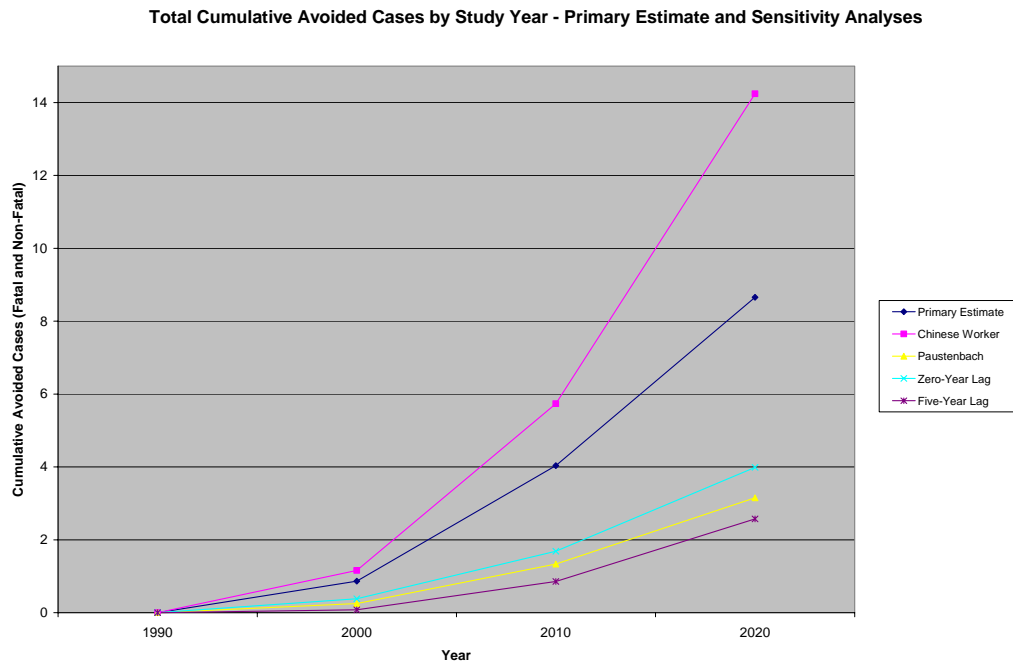
Note: We have linearly interpolated between the avoided leukemia estimates for each target year; however, the true shape of the curve between each of these points is uncertain.

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 significant 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 nine fewer cases of leukemia would occur in the three-county area in the study period, five of which we expect would have been fatal. We estimate the net present value (NPV) in 1990 of the five fatal and four non-fatal leukemia cases avoided to be $13 million in 1999 dollars, using a five percent discount rate. We also expect benefits from the benzene changes that occur between 1990 and 2020 will continue accruing through at least 2030, potentially avoiding another 3 to 4 leukemia cases between 2020 and 2030.

- 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. Inclusion of CAAA-related benzene reductions in the garages of these homes could increase benefits by as much as 20 to 100 percent over 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:

- 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. As a first step, we recommend conducting a reduced-form reanalysis of benefits in the Houston area using NATA concentration data and comparing the results with the findings of this case study. We also recommend review of the regional cancer and non-cancer risk drivers identified in the 1999 NATA to identify high priority HAPs that could potentially be analyzed using this methodology.

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
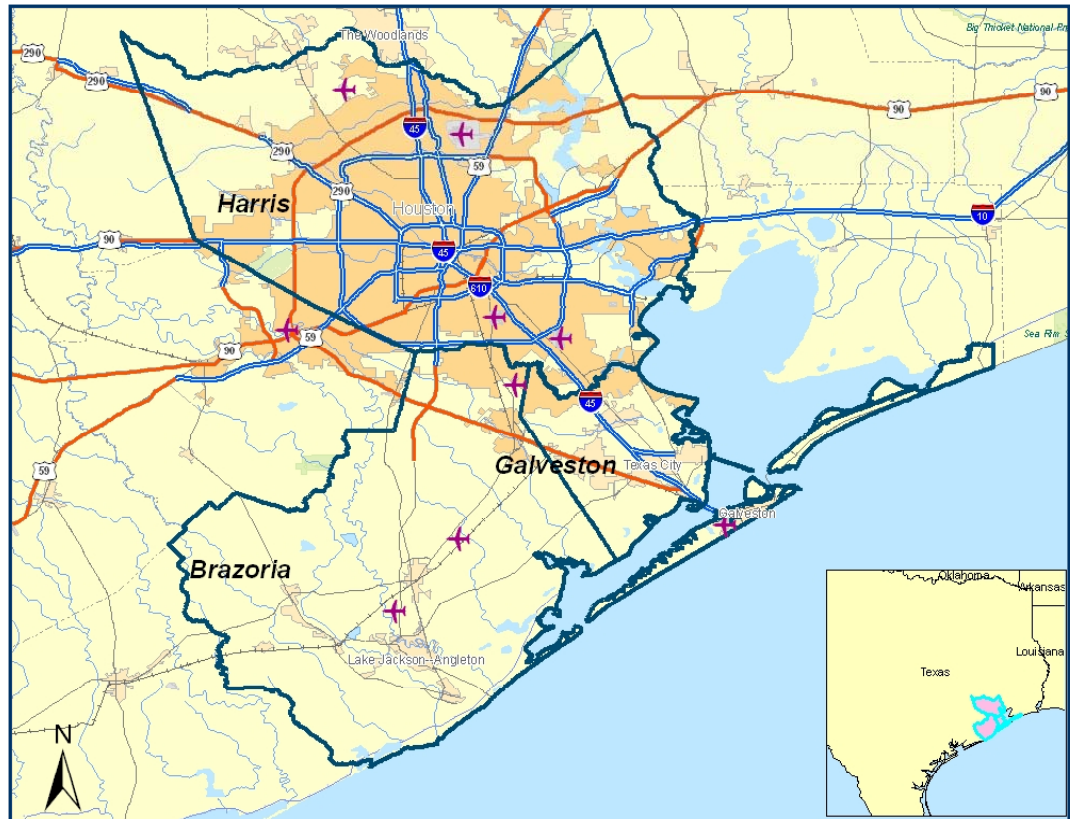
1.2 ORGANIZATION OF THIS DOCUMENT

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 four 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, and Appendix D presents additional details about the health benefits model.

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	**Tailpipe standards** **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 **Evaporative Emissions** Stage II Vapor Recovery Systems (Section 182) Onboard Refueling Vapor Recovery (Section 202; 1998 model year and on) Evaporative Test Procedure **Fuel Regulations** 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	**Maximum Achievable Control Technology (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: 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]

---

[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

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

- 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.

---

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.

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.

To estimate the specific benzene emissions associated with NONROAD's various categories of VOC emissions, we compiled engine-specific benzene speciation factors for

---

[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.

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 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

---

[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.

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.

- **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

---

[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.

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 results using monitor observations obtained from EPA's Air Toxics Archive.[19] We

---

[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."

[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.

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 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

---

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

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 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

---

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

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.

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 $\mu g/m^3$, based on EPA's NATA study).

---

[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).

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]

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

---

[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).

leukemia cases (12) 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

---

[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.

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).

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.

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

---

[29] Only borderline significant results were found for an intensity-dependent nonlinear model, using the Paustenbach exposure estimates.

[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.

[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.

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 latency. See the section entitled "Incorporating Latency" 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 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 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 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 Latency" below for more information.

---

[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).

FIGURE 2: LIFE-TABLE MODEL OVERVIEW
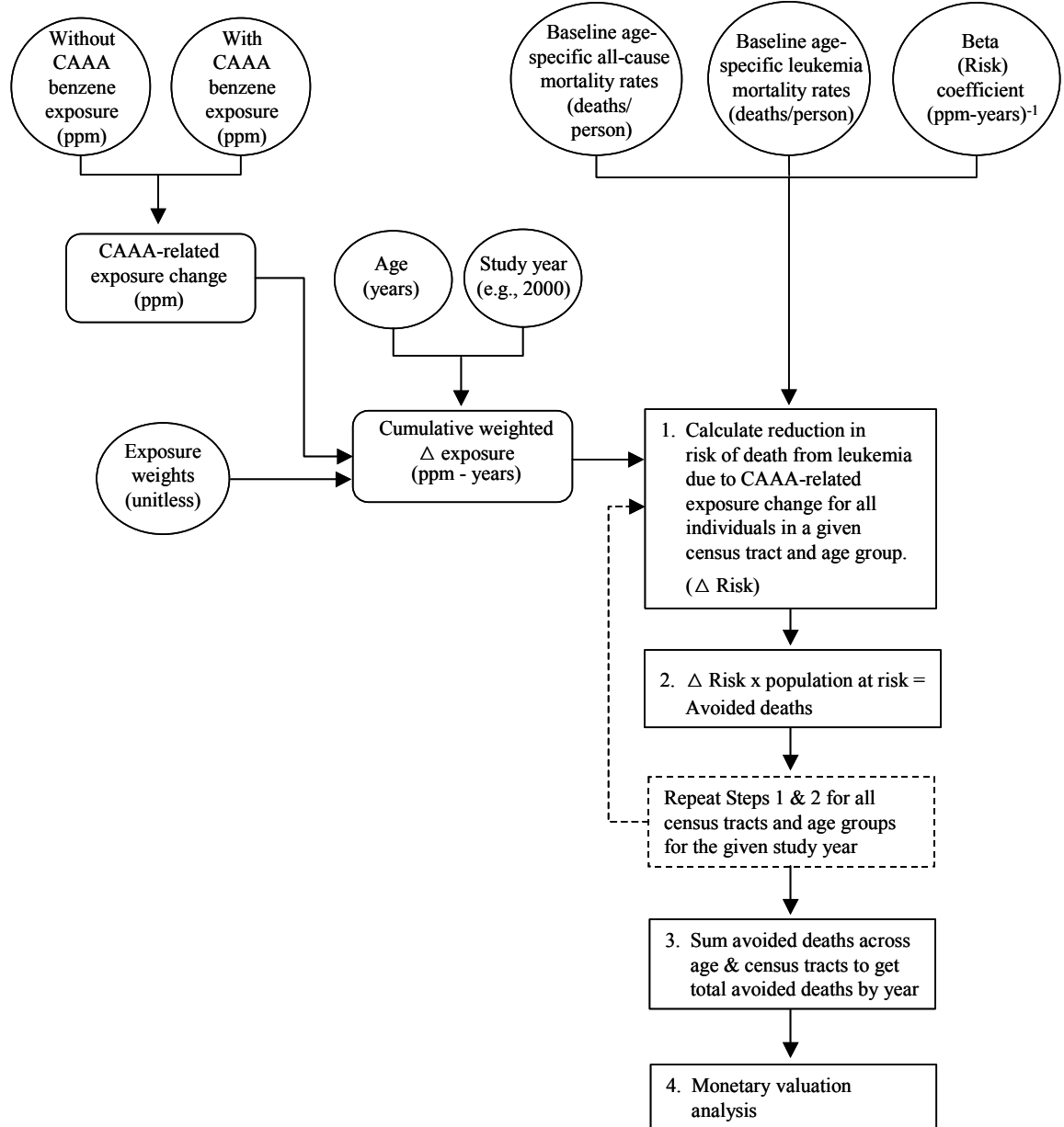
Note: This flowchart assumes the model is being run with leukemia mortality data. The model can also be run with leukemia incidence data. The difference between the model results for these two runs represents an estimate of avoided non-fatal cases of leukemia.

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.)

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

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 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.

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.

Population Data

We used population data from the 2000 US Census.[35] This was available at the census tract level by single year of age. We summed the population at five-year intervals to match the age cohorts 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.[36] 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.[37] 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.[38]

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.[39] 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.[40] In addition, HAPEM6 results were only available for the base year (1990) and target years (2000, 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-

---

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

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

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

[38] 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).

[39] 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).

[40] 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.

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 μg/m³ to ppm.[41] 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.[42]

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 Latency

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 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.[43]

---

[41] In order to convert benzene concentrations from μg/m³ to ppm, we first converted μg/m³ to mg/m³ 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).

[42] The conversion factor consisted of the following: 7/5 days/week $\times$ 24/8 hours/day $\times$ 20/10 m³/day. The value of 10 m³/day is an average breathing rate for the general population. Subjects in the occupational studies on which the risk estimates are based are assumed to have a higher average breathing rate of 20 m³/dayduring an eight-hour work day.

[43] 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).

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 then calculated cumulative avoided deaths for each target year by summing the annual avoided deaths prior to the target year for each age group/census tract combination. Finally, we summed the avoided deaths across all age groups in all census tracts, resulting in total avoided deaths for each county in each target year as well as calculating an overall cumulative sum across the entire study period.

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 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:

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

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

IUR = benzene inhalation unit risk estimate from IRIS $(\mu g/m^3)^{-1}$.[44]

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.

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 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.

---

[44] 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.

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;[45] and

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

---

[45] 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).

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 consistent with those employed to value the benefits of the Second Prospective analysis of criteria pollutants (see Chapter 8). For example, the valuation of fatal cancers relied primarily on the base value of statistical life (VSL) estimates used for particulate matter (PM)-mortality valuation. 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 is 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 the Value of Statistical Life (VSL) estimate presented in the PM National Ambient Air Quality Standards (NAAQS) Regulatory Impact Analysis (RIA) (USEPA, 2006; Table 5-11, page 5-50). We used a median value of $5.5 million in 1990 (in 1999$). 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.. The procedure we applied is described in the PM NAAQS RIA, and results in a VSL value for each year subsequent to 1990 (USEPA, 2006; Table 5-3, page 5-12). The resulting VSL estimate for 2020, for example, was $6.6 million in 1999 dollars.

Pre-Mortality Morbidity

For this analysis, in addition to using EPA's recommended 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.[46]

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).[47] 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 by EPA.[48] 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).[49]

---

[46] See US EPA. (2001a). *Arsenic Rule Benefits Analysis: An SAB Review*. Science Advisory Board. EPA-SAB-EC-01-008, August 30, 2001; and US EPA. (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.

[47] 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.

[48] 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).

[49] The full quote from EPA, 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 individuals willingness-to-pay (WTP) to avoid fatal cancer; hence they cannot be used to justify the proposed

Based on this advice, we conclude that the VSL applied to value avoided fatal leukemia risks represents the value of avoiding a premature death, but 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.[50] This cost can be updated to 1999 price levels using the Consumer Price Index for Medical Care (see USGPO 2006); the result is $108,686, which for our purposes we round to $110,000 and apply as a point estimate to each fatal case of leukemia in the benefits model.[51]

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 recommended that the chronic bronchitis value be supplemented by a value from the one study that values a non-fatal cancer, Magat et al. (1996).[52] The Magat et al. study

---

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).

[50] 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).

[51] CPI-Medical Care series taken from Table B-60 in USGPO 2006, *Economic Report of the President*, Government Printing Office: Washington, DC, transmitted to Congress February 2006.

[52] 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

was 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 adopted the SAB 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. We derived a value for chronic bronchitis from EPA's September 2006 PM NAAQS RIA ($340,000, 1999$, 1990 income levels)[53], and used the VSL from that RIA ($5.5 million, 1999$, 1990 income levels) along with the risk-risk ratio estimated by Magat et al. (0.583) to generate the endpoints of this range.

Developing a central estimate from this range is not straightforward. We evaluated two choices for estimating a central value: the midpoint of the range; and a geometric mean. We prefer the latter value because we believe the values are less likely to be symmetrically distributed and more likely to be skewed in distribution, with a long tail of higher values. The resulting geometric mean is approximately $1 million at 1990 income levels (1999$). We then applied elasticity values for severe and chronic health effects to account for projected real income growth, again following the procedure used in the PM NAAQS RIA, to determine the non-fatal cancer value for each year subsequent to 1990 (USEPA, 2006; Table 5-3, page 5-12). The value we used for 2020 was approximately $1.3 million (1999$).

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 of the direct medical cost of illness as $80,000 (using a 7 percent discount rate) to $150,000 (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.

---

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).

[53] 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.

We also identified a value for a case 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)[54] was based on a U.S. COI study that included indirect costs of illness in the form of lost wages.[55] Converting to 1999$, we obtain a value of $590,000. This estimate is roughly half of the central estimates derived above; a ratio of two is at the low end of estimates of the ratio of WTP to COI for similar health effects, suggesting that our proposed range for valuation of this nonfatal health effect is reasonable.

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.

---

[54] 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.

[55] 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.

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 3 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.[56] 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.[57] 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

---

[56] 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.

[57] 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
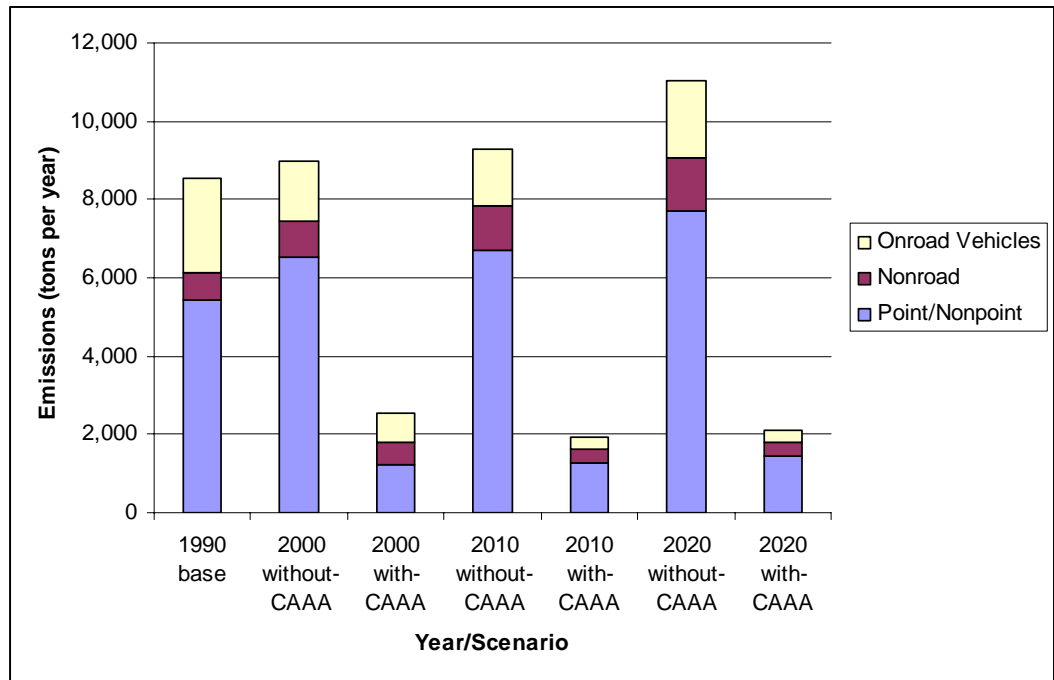
TABLE 3:   HOUSTON-GALVESTON BENZENE EMISSIONS SUMMARY (TONS PER YEAR [TPY])

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.[58] 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.[59] 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

---

[58] 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.

[59] 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 ($\mu g/m^3$) DISTRIBUTIONS FOR 1990, 2000, 2010, AND 2020 FOR *WITH-CAAA* (YELLOW) AND *WITHOUT-CAAA* (GREEN) SCENARIOS
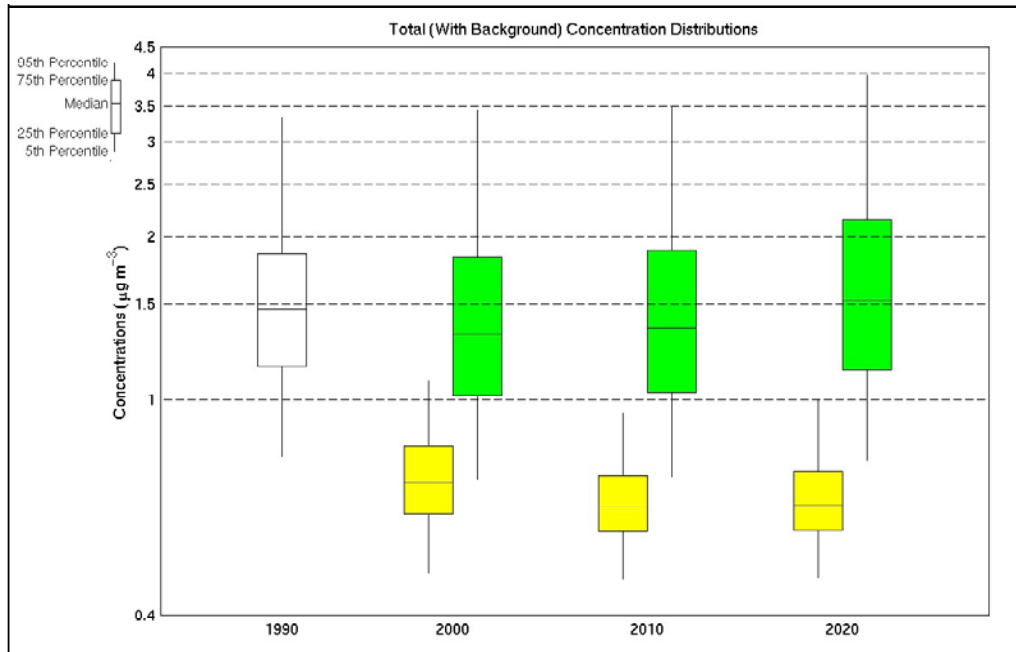
Figure 5 presents maps showing the spatial distribution of benzene reductions across the study area. The top row of maps shows the AERMOD estimates of the reduction in annual average ambient benzene levels due to CAAA programs in (from left to right) 2000, 2010, and 2020. The bottom row shows the same 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 $\mu g/m^3$) 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 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
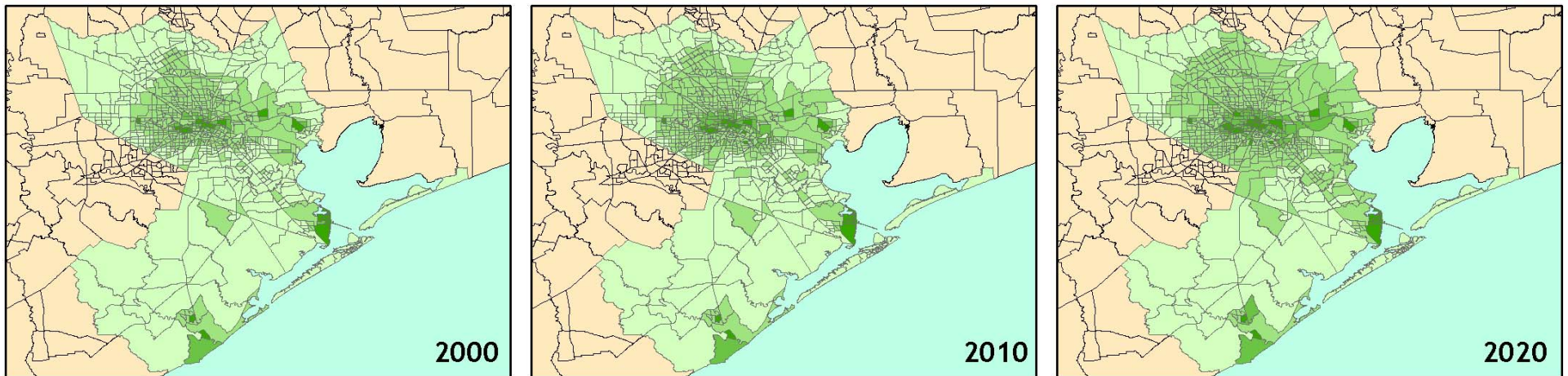
HAPEM RESULTS

Reductions in Concentration   >2.5 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.

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 4 and 5 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 5. That is, the mean estimates in Table 4 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 $\mu g/m^3$, though the range of reductions can be significant, in several cases exceeding 20 $\mu g/m^3$. 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 4:   POPULATION-WEIGHTED MEAN REDUCTION IN AMBIENT ANNUAL AVERAGE BENZENE CONCENTRATION DUE TO CAAA, BY YEAR AND COUNTY

STUDY YEAR	MEAN CHANGE IN BENZENE CONCENTRATION, $\mu g/m^3$ (RANGE)		
	BRAZORIA	GALVESTON	HARRIS
2000	1.0 (0.04 - 25)	0.8 (0.04 - 18)	0.8 (-3 - 34)*
2010	1.0 (0.08 - 25)	0.9 (0.05 - 17)	1.0 (-4 - 33)*
2020	1.2 (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 $\mu g/m^3$ or less. The smallest reductions estimated were between 0.02 and 0.1 $\mu g/m^3$.

TABLE 5:   HAPEM-ESTIMATED MEAN REDUCTION IN ANNUAL BENZENE EXPOSURE CONCENTRATION DUE TO CAAA, BY YEAR AND COUNTY

STUDY YEAR	MEAN CHANGE IN BENZENE CONCENTRATION µg/m³* (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 µg/m³.

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 6 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). Because leukemia is a rare disease, we present the results as the cumulative number of avoided cases, rather than an estimate of annual cases. Therefore, the values in Table 6 represent the total number of expected cases avoided from 1990 through the study year. 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. These additional avoided leukemia cases are not included in the reported cumulative estimate in 2020. We discuss the magnitude of this potential underestimation in benefits further below.

TABLE 6:   CUMULATIVE AVOIDED LEUKEMIA CASES (FATAL AND NON-FATAL) BY STUDY YEAR DUE TO CAAA-RELATED BENZENE EXPOSURE CHANGES IN THE HOUSTON AREA

STUDY YEAR	CUMULATIVE AVOIDED CASES OF LEUKEMIA		
	AVOIDED FATAL CASES	AVOIDED NON-FATAL CASES	TOTAL AVOIDED CASES
1990	0	0	0
2000	0.5	0.4	0.9
2010	2	2	4
2020	5	4	9

Our results indicate that by the year 2020, a total of nine cases of leukemia would be avoided due to the 1990 CAAA programs in the Houston area, with seven of those occurring in Harris County, and approximately one each in Brazoria and Galveston Counties. We estimate five of the nine cases to be fatal and four to be non-fatal.

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 7.

TABLE 7:   TOTAL CUMULATIVE MONETARY BENEFITS THROUGH 2020 DUE TO CAAA-RELATED CHANGES IN BENZENE EXPOSURE IN THE HOUSTON AREA

DISCOUNT PERCENTAGE	TOTAL BENEFITS (1990 NPV, MILLIONS OF1999$)		
	BENEFITS FROM FATAL CASES OF LEUKEMIA	BENEFITS FROM NON-FATAL CASES OF LEUKEMIA	TOTAL BENEFITS
Primary Estimate (5%)	$11	$2	$13
No Discounting	$29	$5	$34
Low Discount Rate (3%)	$16	$3	$19
High Discount Rate (7%)	$8	$1	$9

The values in Table 7 represent the total net present value estimate (discounted to 1990) of the benefits of the CAAA-related benzene controls in Houston between 1990 and 2020.[60] Our primary estimate of benefits due to CAAA-related reductions in benzene are $13 million (in 1999$), $11 million of which are due to fatal cases of leukemia, and $2 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. 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 also presented in Table 7 and range from $9 million for the high discount rate to $34 million when no discount rate is applied.

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 and are not captured in the 2020 cumulative avoided cases value in Table 6.

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

---

[60] 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.

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 one additional fatal case and one additional non-fatal case avoided in the first five years after the study period, yielding a total of 11 avoided cases of leukemia. By the year 2030, we estimated one additional fatal and 0.6 of an additional 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 12.5.

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 8 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 8, 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.[61] 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

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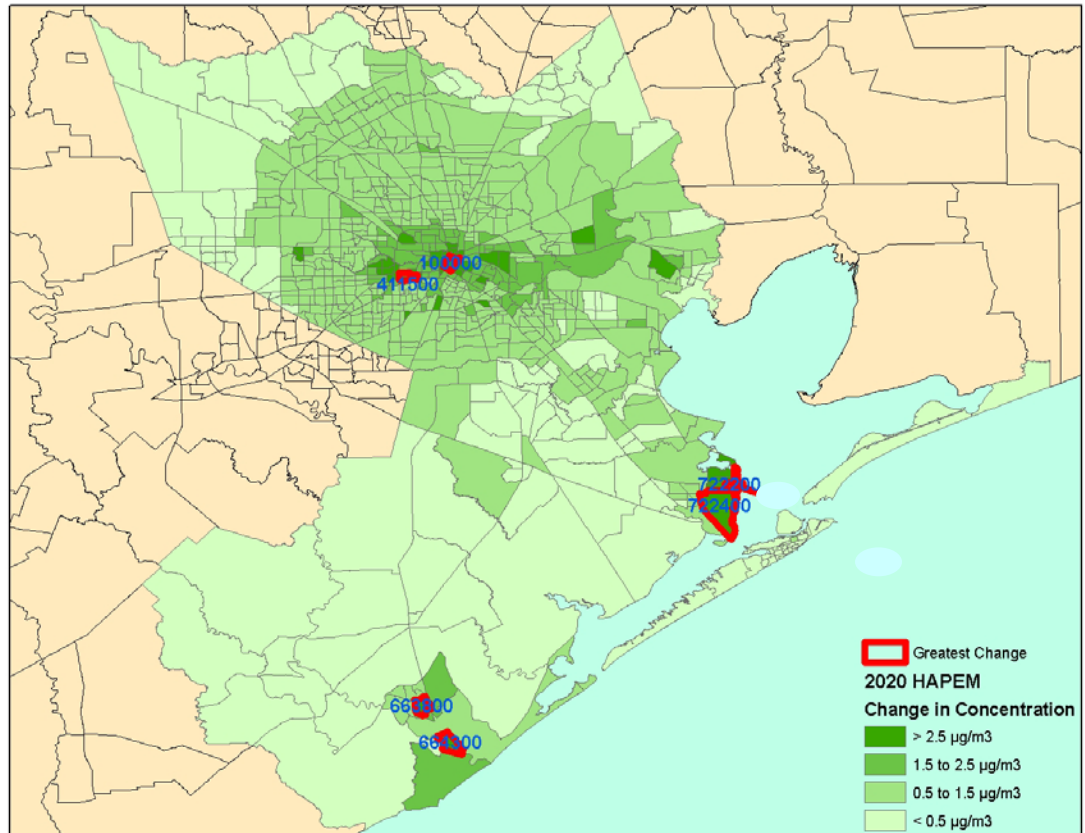
[61] Risks were calculated using the 7.8×10^{-6} per μg/m^3 benzene inhalation unit risk (IUR) from the range of IURs reported on IRIS.

TABLE 8:   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 $\mu g/m^3$ 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
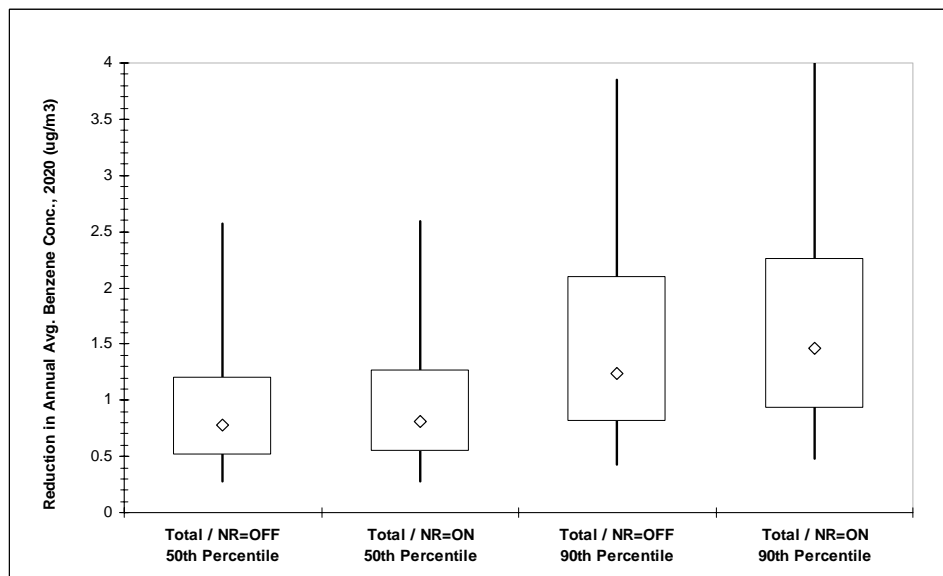
TABLE 9:   CAAA-RELATED BENZENE REDUCTIONS IN 2020 INDIVIDUALS LIVING IN CENSUS TRACTS WITH HIGH AMBIENT BENZENE CONCENTRATIONS DUE TO ON-ROAD SOURCES - EFFECT OF HAPEM NEAR-ROADWAY ALGORITHM

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 89 – 90 percent. If the average exposure estimate attributable to attached garages (1.2 μg/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 μg/m^3. This would correspond to an additional 0.1 – 0.5 annual avoided cases of leukemia in the Houston area in 2020.[62] Results of this magnitude suggest that adding attached garage-related benefits to our primary estimate of annual avoided cases of leukemia in 2020 (0.5 cases) could increase benefits by as much as 20 to 100 percent over our primary estimate.[63]

3.3.4 SENSITIVITY ANALYSES

We performed four 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.[64] We first 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. We then 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. Finally, we ran the model with two alternate lags, a zero-year lag and a five-year lag.

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.

---

[62] Note that the low-end estimate relied on the lower range of percentages of emissions likely to occur within attached garages as well as the lower bound of the range of IUR values presented in IRIS.

[63] 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.

[64] 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

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).[65] 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 14 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.[66] 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). 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 six avoided cases of AML (fatal and non-fatal) between 1990 and 2020 due to CAAA-related changes in benzene exposure.   We then compared national AML mortality rates to AML incidence rates and found that 60 percent of incident cases were fatal. Therefore, we would expect that of the six avoided cases of AML, four would be fatal and two 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,

---

[65] 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, 2005) 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.

[66] 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/.

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.[67] 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 three 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 HAPEM6.[68] 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 four avoided cases between 1990 and 2020 and the five-year lag yielded an estimate of three.

Summary

Table 10 displays the primary estimate of cumulative avoided cases (fatal and non-fatal) of leukemia by study year as well as estimates for the sensitivity analyses. Total avoided cases between 1990 and 2020 for the primary estimate is nine and the sensitivity analyses range between three and 14. Figure 8 presents the total cumulative cases of leukemia between 1990 and 2020 for the primary case as well as three of the sensitivity analyses in graphical form.

---

[67] 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.

[68] 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.

TABLE 10:   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	CHINESE WORKER COHORT	PAUSTENBACH EXPOSURE MATRIX	ZERO-YEAR LAG	FIVE-YEAR LAG
1990	0	0	0	0	0
2000	0.9	1	0.2	0.4	0.08
2010	4	6	1	2	0.9
2020	9	14	3	4	3

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

Total Cumulative Avoided Cases by Study Year - Primary Estimate and Sensitivity Analyses
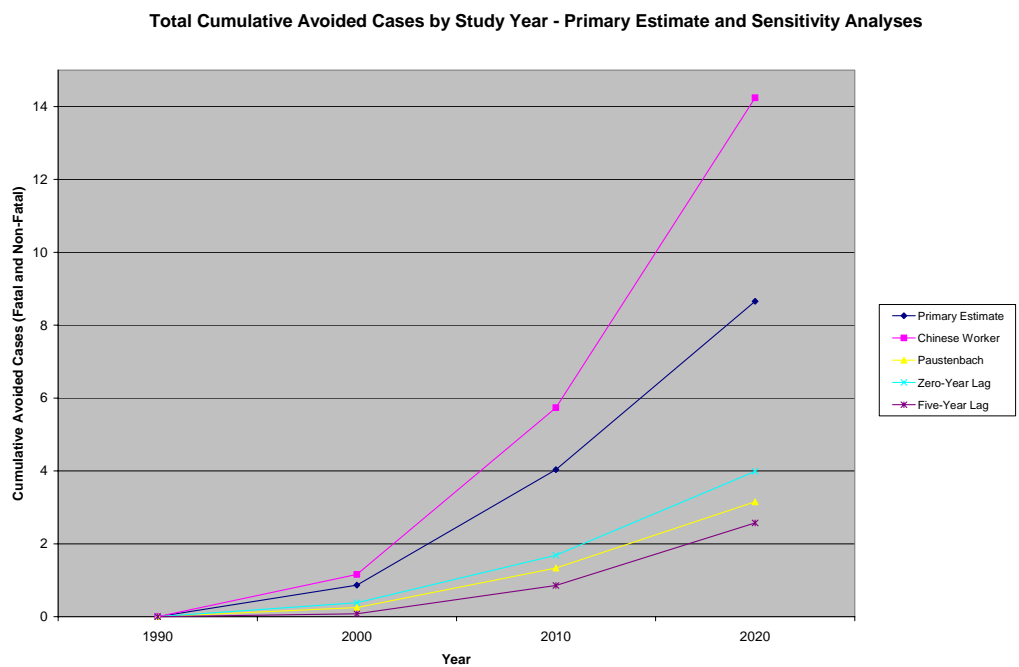
Note: We have linearly interpolated between the avoided leukemia estimates for each target year; however, the true shape of the curve between each of these points is uncertain.

We also assessed the economic benefits associated with the avoided cases of leukemia for the sensitivity analyses. Table 11 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 MONETARY BENEFITS OF CAAA-RELATED REDUCTIONS IN BENZENE IN THE HOUSTON AREA – PRIMARY ESTIMATE AND SENSITIVITY ANALYSES RESULTS (IN MILLIONS OF 1999$)

PRIMARY ESTIMATE	CHINESE WORKER	PAUSTENBACH	ZERO-YEAR LAG	FIVE-YEAR LAG
$13	$19	$4	$5	$3

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 significant 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 nine fewer cases of leukemia would occur in the three-county area in the study period, five of which we expect would have been fatal. We estimate the net present value (NPV) in 1990 of the five fatal and four non-fatal leukemia cases avoided to be $13 million in 1999 dollars, using a five percent discount rate. We also expect benefits from the benzene changes that occur between 1990 and 2020 will continue accruing through at least 2030, potentially avoiding another 3 to 4 leukemia cases between 2020 and 2030.

- 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. Inclusion of CAAA-related benzene reductions in the garages of these homes could increase benefits by as much as 20 to 100 percent over 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 UNCERTAINTY

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. The uncertainties likely to have the greatest impact on our results 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 an ME for attached garages in the exposure modeling step; potential underestimation of benzene levels by AERMOD; and emissions uncertainties related to modeling on-road and non-road sources. We believe that overall, the uncertainties in our analysis are likely to cause our results to be underestimated. We describe the potential uncertainties of the study in greater detail below.

4.2.1 EMISSIONS UNCERTAINTY

In general, the major potential sources of emissions uncertainty are expected to be similar to other benefits assessment studies; that is, the selection of a base emissions inventory and the application of growth and control factors to the base estimates to project future year emissions under each scenario. Emissions inventories can vary in the methods and underlying assumptions used to estimate emissions. Different years of the same inventory (e.g., 1999 vs. 2002 NEI) can vary as well due to refinements in the emissions estimation process. While it was not possible to run our analysis using alternative inventories, we did compare our chosen inventories with potential alternatives. Comparisons of base year inventories for stationary and non-road sources with alternative inventories, such as the 1999 NEI and 2002 Toxics Release Inventory (TRI), showed good agreement (Pechan, 2006 and Thurman, 2006). Comparison of 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 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. However, 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.

Some categories of emissions were not included in our analysis, likely biasing our results low. These 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). In addition, emission controls from the MSAT program, which was established after the *With-CAAA* scenario was fixed, are not included. 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.

We acknowledge additional uncertainty in predicting future growth rates and control factors. We applied the Annual Energy Outlook (AEO) 2005 growth rates for consistency with the main 812 analysis; however, unlike that analysis, we did not perform alternative high- and low-growth scenarios due to resource limitations. To the extent that the AEO growth predictions under- or over-estimate future growth, our estimate of the difference in emissions due to the CAAA will likely be under- or over-estimated. Similarly, errors in assumptions about compliance and/or control efficiencies could lead us to mis-estimate benzene emissions reductions as well.

4.2.2 AIR QUALITY AND EXPOSURE MODELING UNCERTAINTY

A list of potential sources of uncertainty associated with air quality and exposure modeling can be found in Appendix B. Additional potential issues include:

- 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.[69]

---

[69] 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.

IEc

- 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* 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.

- The HAPEM6 model results are associated with some uncertainty. For instance, HAPEM6 calculates long-term average exposure concentrations and therefore does not preserve the time-sequence of exposure events when sampling from the time/activity databases. In addition, some of the MEs used in HAPEM6 were derived from measurement or modeling studies in specific locations. These studies are subject to measurement and modeling uncertainties and also may not be representative of other locations (ICF, 2007).

- Using modeled exposure concentrations, as opposed to measured, involves some uncertainty. Actual personal exposures could be higher or lower than modeled concentrations.

- 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 found the annual avoided cases of leukemia estimated in our primary estimate could potentially be increased by 20 to 100 percent in 2020.

- Our analysis assumed a constant background benzene concentration to account for transport of benzene into the study area from surrounding counties. We assumed no changes to the background benzene contribution in future years under the *With-* or *Without-CAAA* scenarios. However, CAAA regulations are likely to have contributed to some reductions in benzene transported across county lines when compared with the counterfactual scenario; thus our background modeling approach likely biases our benefits estimates downward. The magnitude of this bias is expected to be small, given that winds come primarily from the Gulf of Mexico to southeast (see Appendix B) and that within our study area, concentrations tend to decrease fairly quickly downwind of major point sources.

4.2.3 HEALTH BENEFITS MODELING AND VALUATION

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

- Our model is sensitive to uncertainty in the selection of the dose-response and cessation lag model for benzene-induced leukemia. Sensitivity analyses show that our results can vary by plus 50 percent to minus 67 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.[70]

- 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.[71] Therefore, our results do not provide a comprehensive estimate of health benefits from benzene reductions in the Houston area.

- 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.

- 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.[72]

- 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.

- The leukemia cohort studies are based on occupational exposure levels. Extrapolation of the dose-response function to ambient environmental levels

---

[70] We did not assess the uncertainty in the individual dose-response estimates in this case study. Future analyses could potentially run the model using the bounds of the 95 percent confidence interval around the dose-response slope factors.

[71] These health endpoints could potentially be revisited in future benzene benefits analyses to determine if recent literature developments support quantification of benefits.

[72] 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, 2005) for chemicals with a mutagenic MOA. Early-life adjustments could be explored in a future case study.

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 D, there is some evidence to suggest the function may be supra-linear; if so, we will have underestimated the benefits of CAAA benzene reductions.

- 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 D, 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.

- We based our avoided incidence estimates on population data from the 2000 census. Because we did not project population growth for 2010 and 2020, we have likely underestimated the avoided cases in these future years.[73] In addition, the use of a static population does not incorporate changes in the age structure. The effect of this exclusion would depend on how the age structure in the Houston area is expected to change over time.

- 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.[74]

- The VSL value we applied ($5.5 million in 1999$) 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").

- 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 is derived from 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

---

[73] In fact, one independent projection by Woods and Poole indicates that the population of the Houston Metropolitan Area will increase by roughly 40 percent between 2000 and 2020 (http://www.houston.org/blackfenders/09AW005.pdf).

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

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.[75] 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.[76]

---

[75] 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.

[76] 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.

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.[77]

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.

We also recommend conducting a simplified "back-of-the-envelope" analysis of the benefits of CAAA-related benzene reductions in Harris, Brazoria, and Galveston counties using benzene concentrations from the 1999 NATA and/or the forthcoming 2002 NATA for the *With-CAAA* scenario in 2000. Comparing those results with the findings of this case study will help EPA better assess the value of the more data-intensive urban-scale approach applied in this case study and help them evaluate the usefulness of the NATA for a national-scale HAP benefits assessment.

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. As a first step, we recommend conducting a reduced-form reanalysis of benefits in the Houston area using NATA concentration data and comparing the results with the findings of this case study. We also recommend review of the regional cancer and non-cancer risk drivers identified in the 1999 NATA to identify high priority HAPs that could potentially be analyzed using this methodology.

---

[77] 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.

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.

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.

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.

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). Hematotoxity in workers exposed to low levels of benzene. *Science* 306: 1774-1776.

National Research Council. (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.

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.

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.

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.

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

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.

IEc

APPENDIX A:

BENZENE EMISSIONS MODELING REPORT

IEc

APPENDIX B:

BENZENE AIR QUALITY AND EXPOSURE MODELING REPORT

IEc

APPENDIX C:

BENZENE HEALTH EFFECTS LITERATURE REVIEW MEMO

IEc

APPENDIX D:

LIFE-TABLE MODEL EQUATIONS

IEc

APPENDIX E:

ATTACHED GARAGE ANALYSIS EQUATIONS

APPENDIX B

Air Quality and Exposure Modeling in Support of Section 812 Benzene
Case Study: Methodology and Results

The information in this document has been reviewed in accordance with the U.S. EPA administrative review policies and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for their use.

The following trademarks appear in this document:

UNIX is a registered trademark of AT&T Bell Laboratories.

Linux is a registered trademark of Red Hat

SAS® is a registered trademark of SAS Institute

SUN is a registered trademark of Sun Microsystems, Inc.

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.

9

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 $Factor_{I,X}$ is the allocation factor for airport I, for X type (general aviation, commercial aircraft, air taxi, or military aircraft), $IT_{X,I}$ is the itinerant data for airport I for type X, and $IT_{X,county}$ 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'

15

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

17

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.

19

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

25

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 (3.63×10^{-1})	Galveston (3.97×10^{-1})	Harris (4.64×10^{-1})
1990	Base	Stationary	8.08×10^{-1}	7.27×10^{0}	6.97×10^{-1}
		Mobile	8.90×10^{-2}	2.21×10^{-1}	4.15×10^{-1}
		Total	1.26×10^{0}	7.89×10^{0}	1.58×10^{0}
2000	CAA	Stationary	1.04×10^{-1}	1.35×10^{-1}	1.37×10^{-1}
		Mobile	3.63×10^{-2}	6.60×10^{-2}	1.82×10^{-1}
		Total	5.04×10^{-1}	5.98×10^{-1}	7.83×10^{-1}
	NON- CAA	Stationary	1.40×10^{0}	1.16×10^{0}	9.10×10^{-1}
		Mobile	6.25×10^{-2}	1.11×10^{-1}	3.46×10^{-1}
		Total	1.83×10^{0}	1.66×10^{0}	1.72×10^{0}
2010	CAA	Stationary	1.15×10^{-1}	1.32×10^{-1}	1.42×10^{-1}
		Mobile	1.73×10^{-2}	3.19×10^{-2}	9.56×10^{-2}
		Total	4.95×10^{-1}	5.62×10^{-1}	7.02×10^{-1}
	NON- CAA	Stationary	1.39×10^{0}	1.17×10^{0}	9.19×10^{-1}
		Mobile	6.31×10^{-2}	1.04×10^{-1}	3.70×10^{-1}
		Total	1.81×10^{0}	1.67×10^{0}	1.75×10^{0}
2020	CAA	Stationary	1.28×10^{-1}	1.48×10^{-1}	1.62×10^{-1}
		Mobile	1.55×10^{-2}	2.82×10^{-2}	9.11×10^{-2}
		Total	5.07×10^{-1}	5.73×10^{-1}	7.17×10^{-1}
	NON- CAA	Stationary	1.58×10^{0}	1.34×10^{0}	1.05×10^{0}
		Mobile	7.79×10^{-2}	1.24×10^{-1}	4.71×10^{-1}
		Total	2.02×10^{0}	1.86×10^{0}	1.99×10^{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	County Galveston	County Harris
1990	Base	Stationary	7.32×10^{-1}	6.32×10^{0}	5.67×10^{-1}
		Mobile	1.24×10^{-1}	2.29×10^{-1}	4.71×10^{-1}
		Background	2.93×10^{-1}	3.05×10^{-1}	3.74×10^{-1}
		Total	1.15×10^{0}	6.85×10^{0}	1.41×10^{0}
2000	CAA	Stationary	7.52×10^{-2}	1.03×10^{-1}	1.14×10^{-1}
		Mobile	5.05×10^{-2}	7.59×10^{-2}	2.04×10^{-1}
		Background	2.97×10^{-1}	3.31×10^{-1}	3.77×10^{-1}
		Total	4.23×10^{-1}	5.10×10^{-1}	6.95×10^{-1}
	NON-CAA	Stationary	9.74×10^{-1}	7.69×10^{-1}	7.86×10^{-1}
		Mobile	7.71×10^{-2}	1.21×10^{-1}	3.71×10^{-1}
		Background	2.54×10^{-1}	3.09×10^{-1}	3.53×10^{-1}
		Total	1.31×10^{0}	1.20×10^{0}	1.51×10^{0}
2010	CAA	Stationary	8.05×10^{-2}	1.02×10^{-1}	1.18×10^{-1}
		Mobile	2.45×10^{-2}	3.66×10^{-2}	1.05×10^{-1}
		Background	3.01×10^{-1}	3.34×10^{-1}	3.84×10^{-1}
		Total	4.06×10^{-1}	4.73×10^{-1}	6.07×10^{-1}
	NON-CAA	Stationary	9.62×10^{-1}	7.79×10^{-1}	7.88×10^{-1}
		Mobile	7.86×10^{-2}	1.16×10^{-1}	3.86×10^{-1}
		Background	2.55×10^{-1}	3.10×10^{-1}	3.53×10^{-1}
		Total	1.30×10^{0}	1.20×10^{0}	1.53×10^{0}
2020	CAA	Stationary	8.89×10^{-2}	1.13×10^{-1}	1.35×10^{-1}
		Mobile	2.27×10^{-2}	3.31×10^{-2}	1.01×10^{-1}
		Background	3.00×10^{-1}	3.34×10^{-1}	3.83×10^{-1}
		Total	4.12×10^{-1}	4.80×10^{-1}	6.19×10^{-1}
	NON-CAA	Stationary	1.09×10^{0}	8.82×10^{-1}	8.97×10^{-1}
		Mobile	9.78×10^{-2}	1.38×10^{-1}	4.97×10^{-1}
		Background	2.53×10^{-1}	3.08×10^{-1}	3.51×10^{-1}
		Total	1.44×10^{0}	1.33×10^{0}	1.75×10^{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
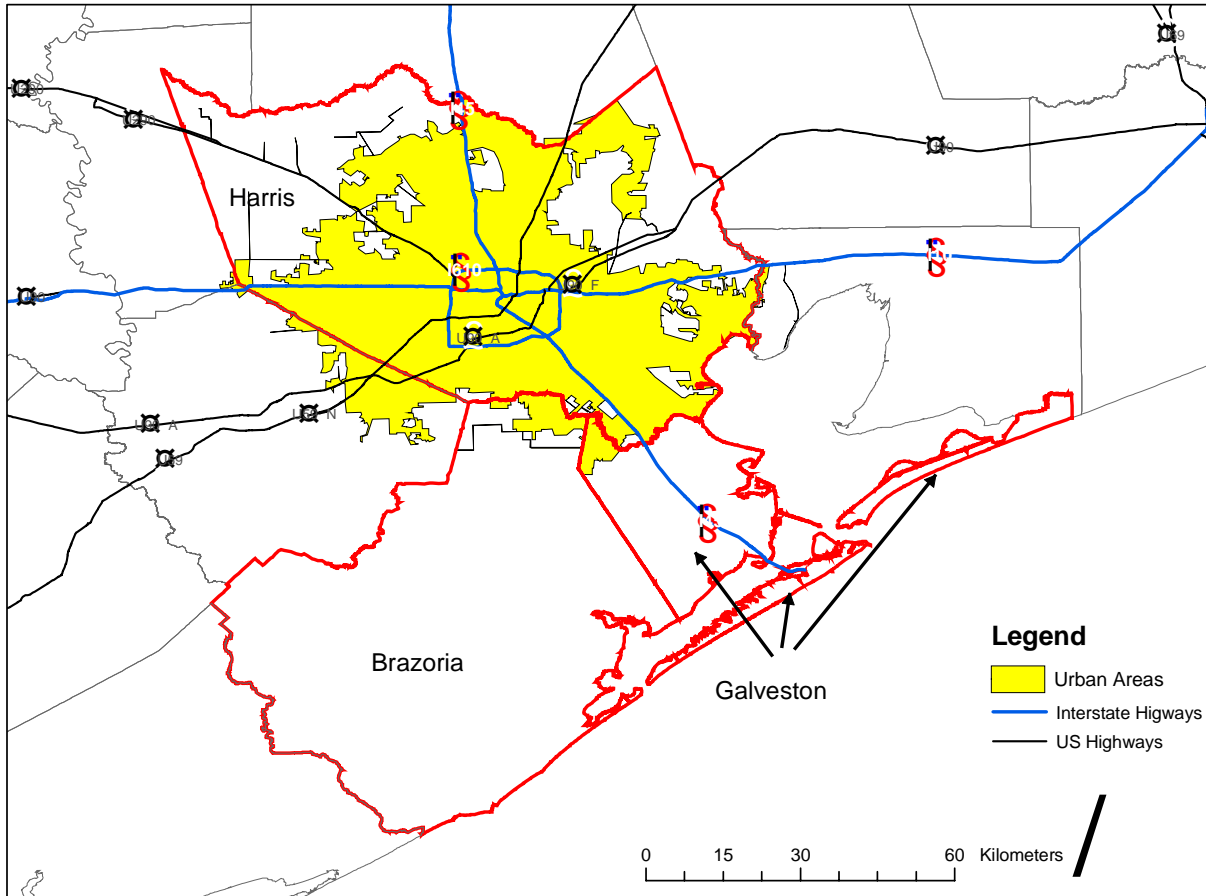
Figure 1. Domain of modeling study. The three counties of emphasis are labeled and outlined in red and the Houston metropolitan area is shown in yellow. Key roads are also shown.
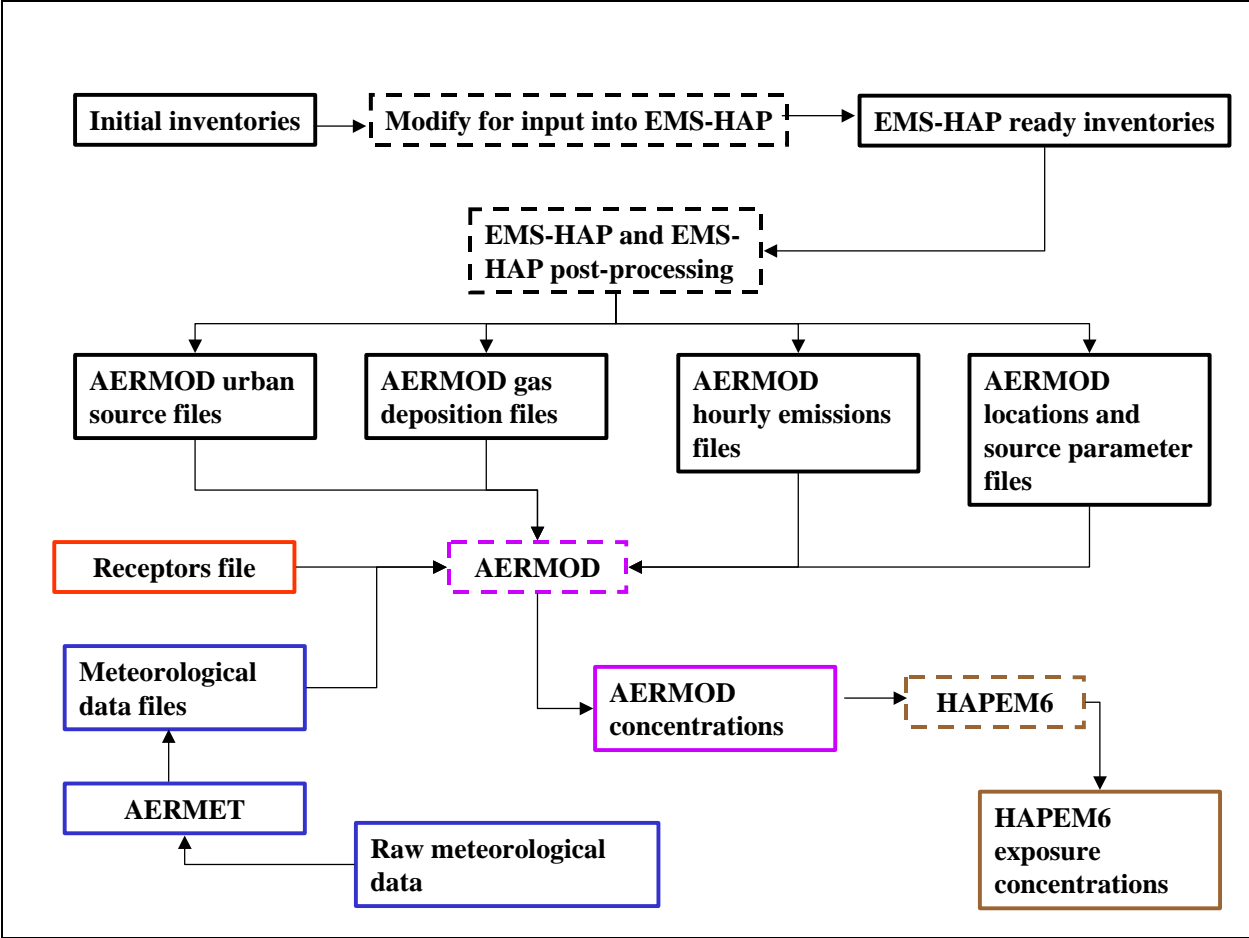
Figure 2. General overview of processing steps from emissions to HAPEM6 exposure concentrations.
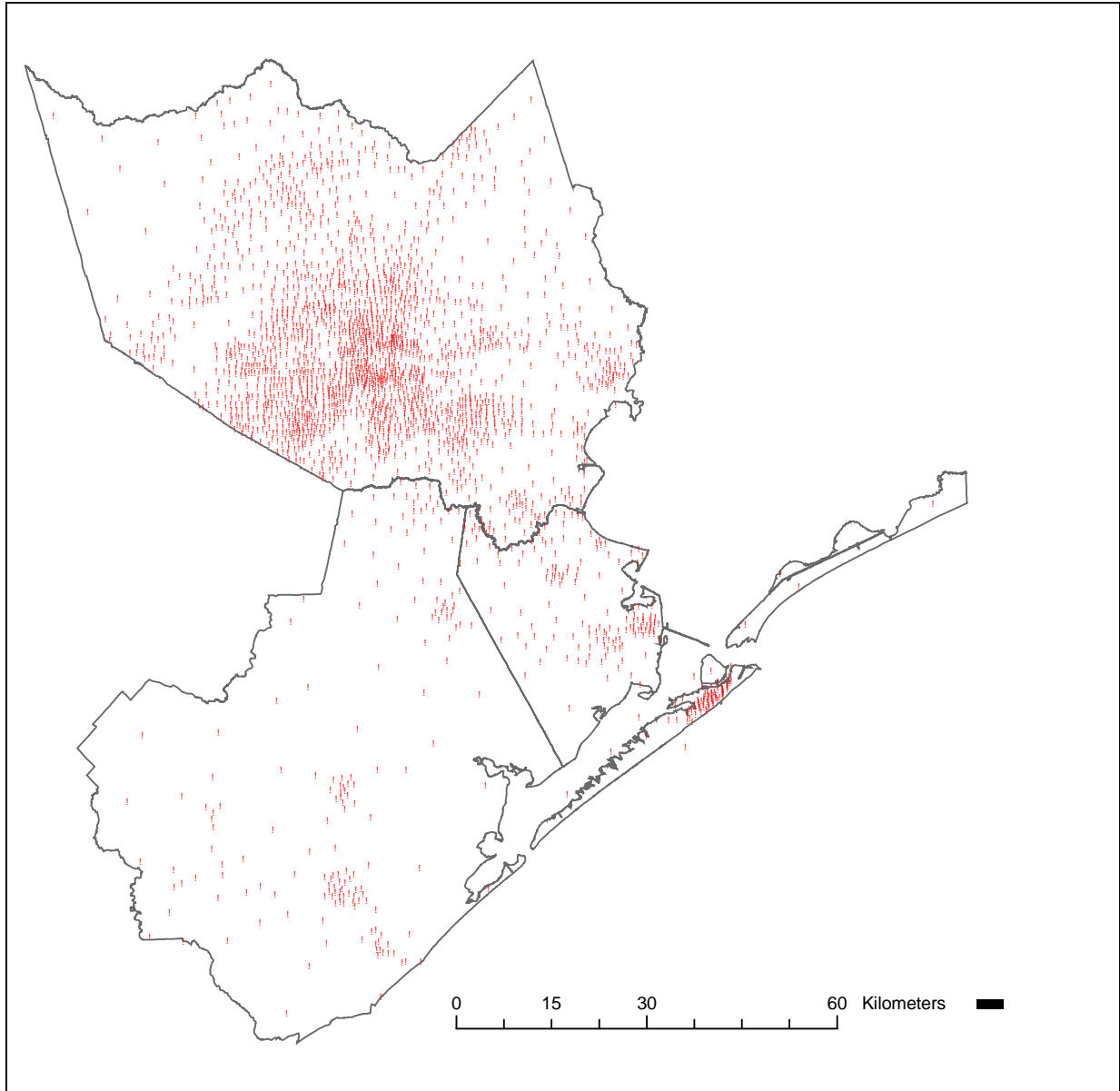
Figure 3. Receptors for 1990 simulation based on 1990 census block group centroids.
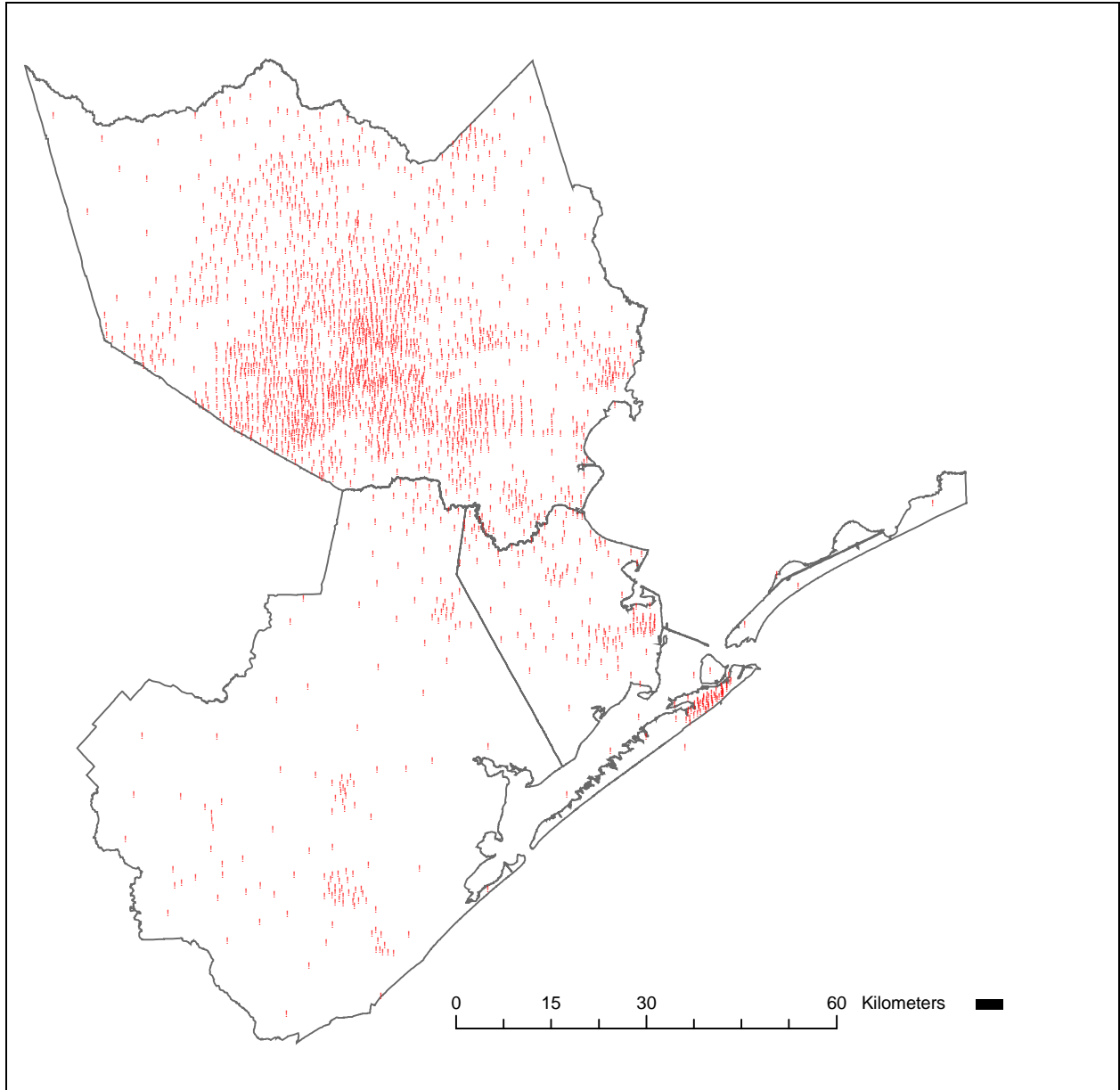
Figure 4. Receptors for simulations of the years 2000, 2010, and 2020 based on 2000 census block group centroids.
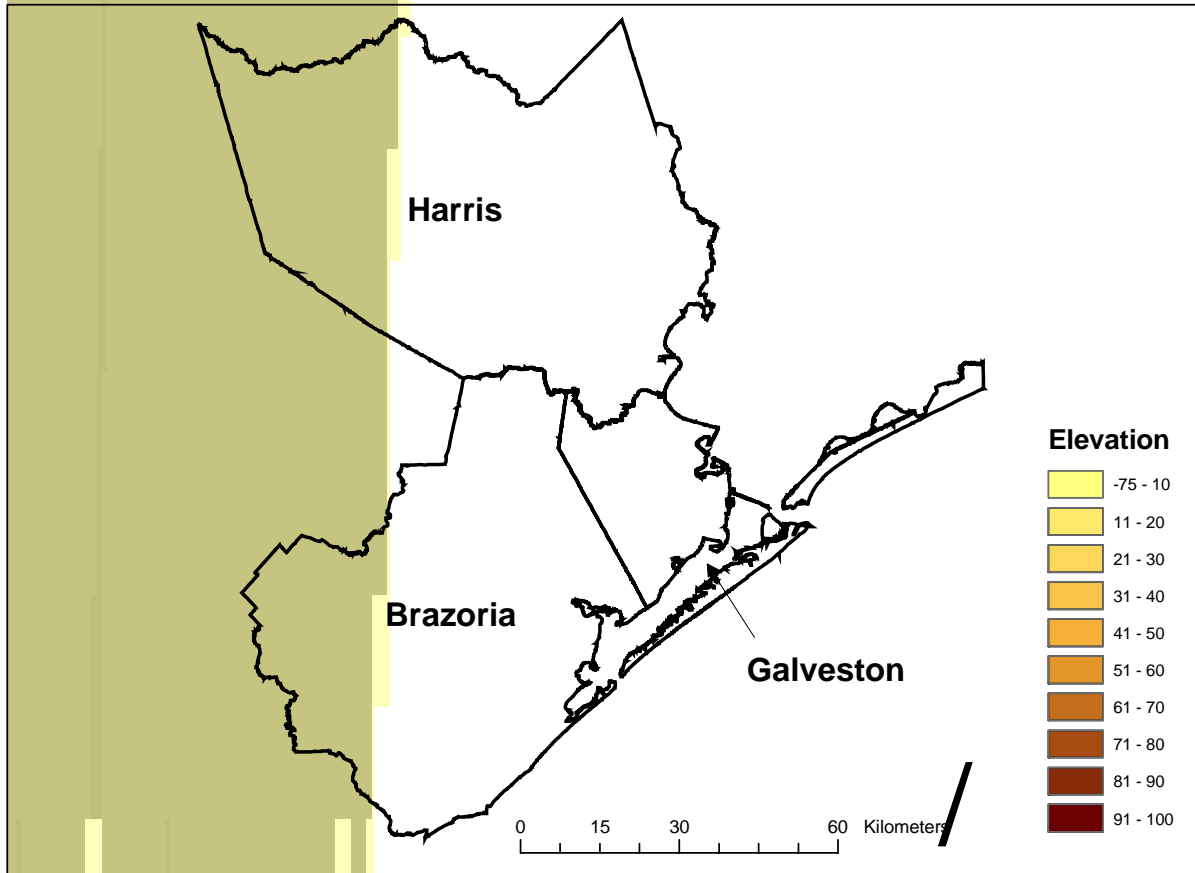
Figure 5. Elevation (m) of the Houston domain.
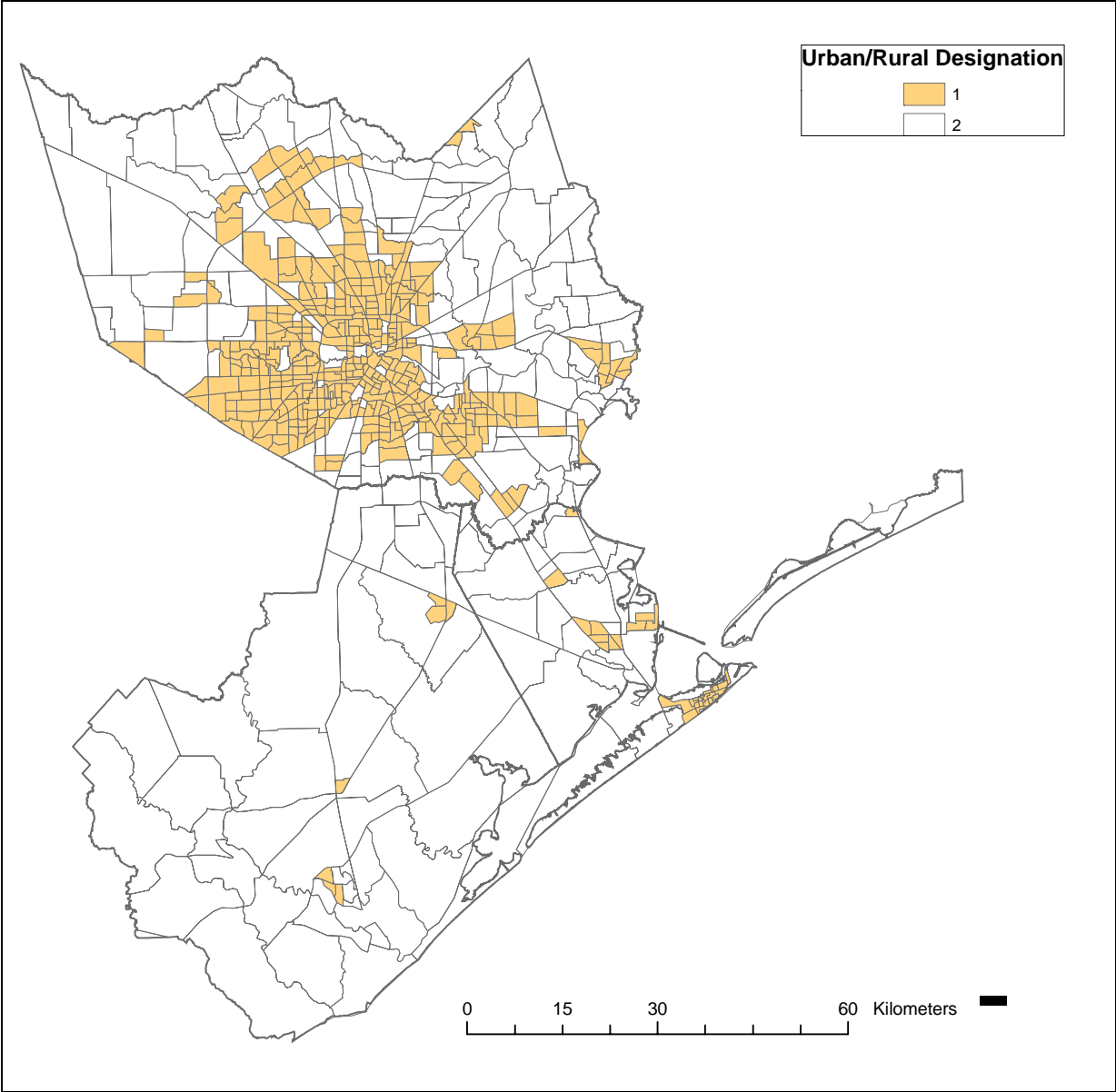
Figure 6. 1990 urban (brown) and rural (white) census tracts.
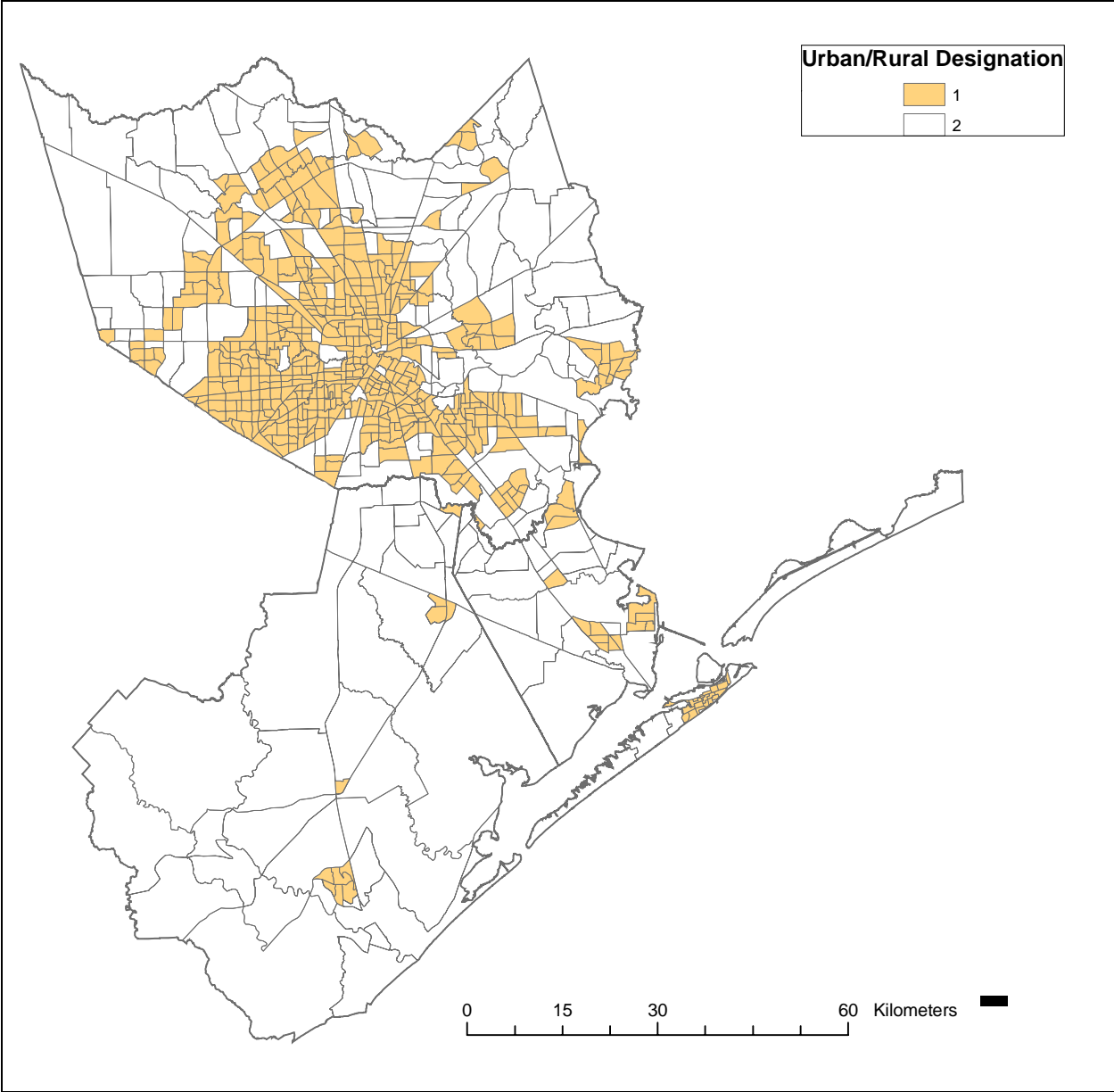
Figure 7. Year 2000 urban (brown) and rural (white) census tracts.
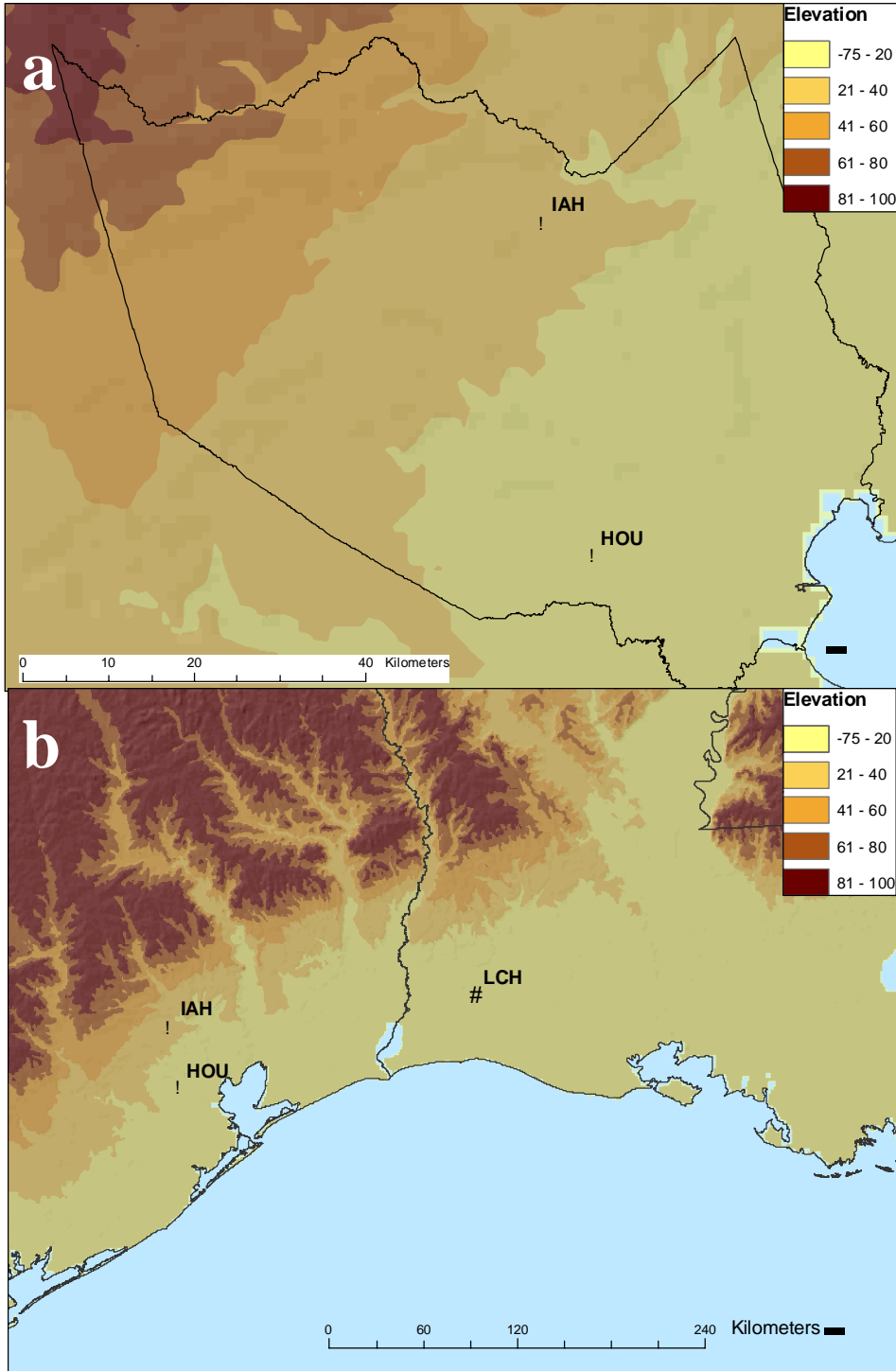
Figure 8. Locations of a) surface meteorological stations for 1990 and 2000 and b) relative locations of surface and upper air stations.

(IAH = George Bush Intercontinental Airport; HOU = Hobby Field; and LCH = Lake Charles)
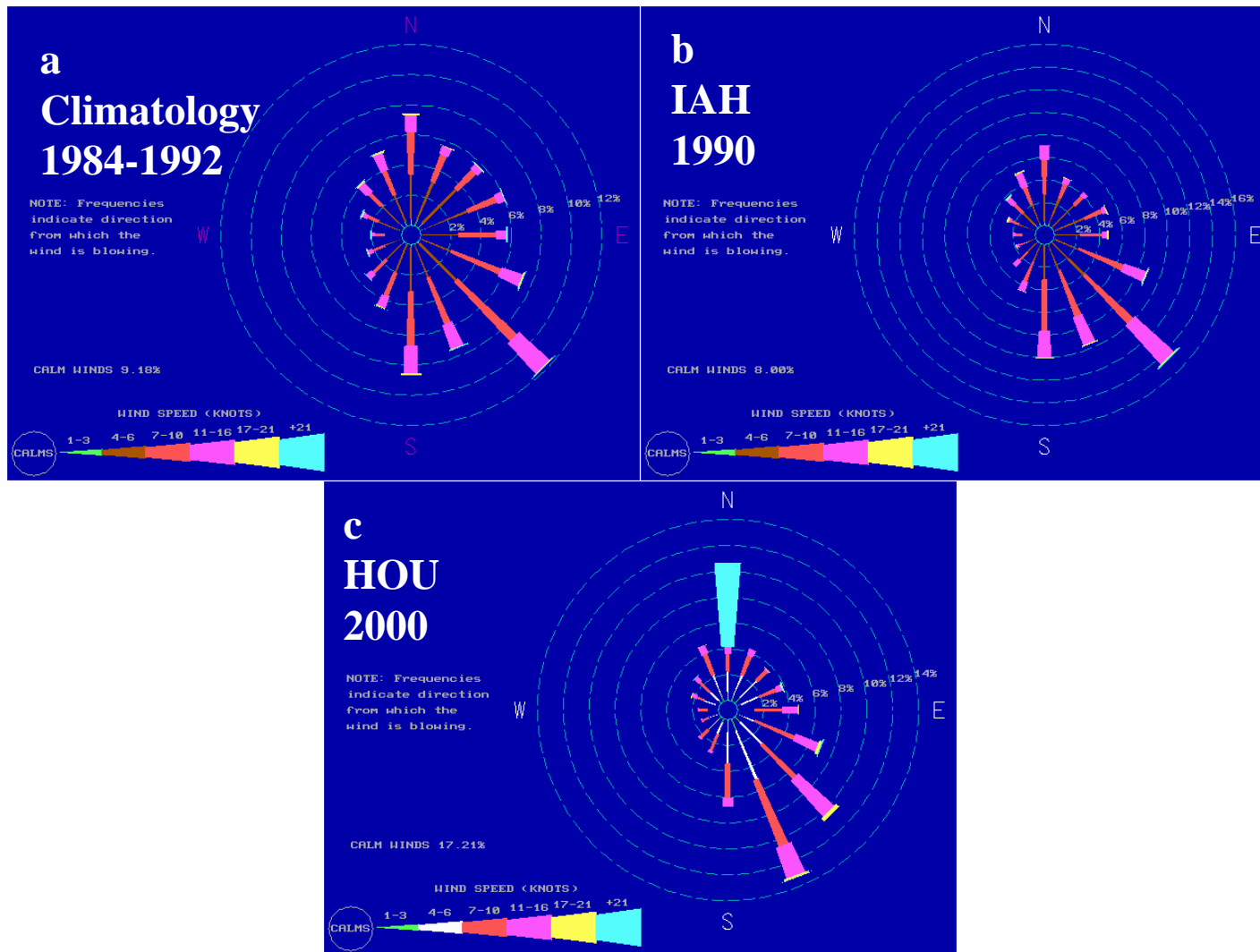
Figure 9. Wind roses for a) 8 year climatology for IAH, b) 1990 IAH, and c) HOU 2000. Winds are generally from the southeast.
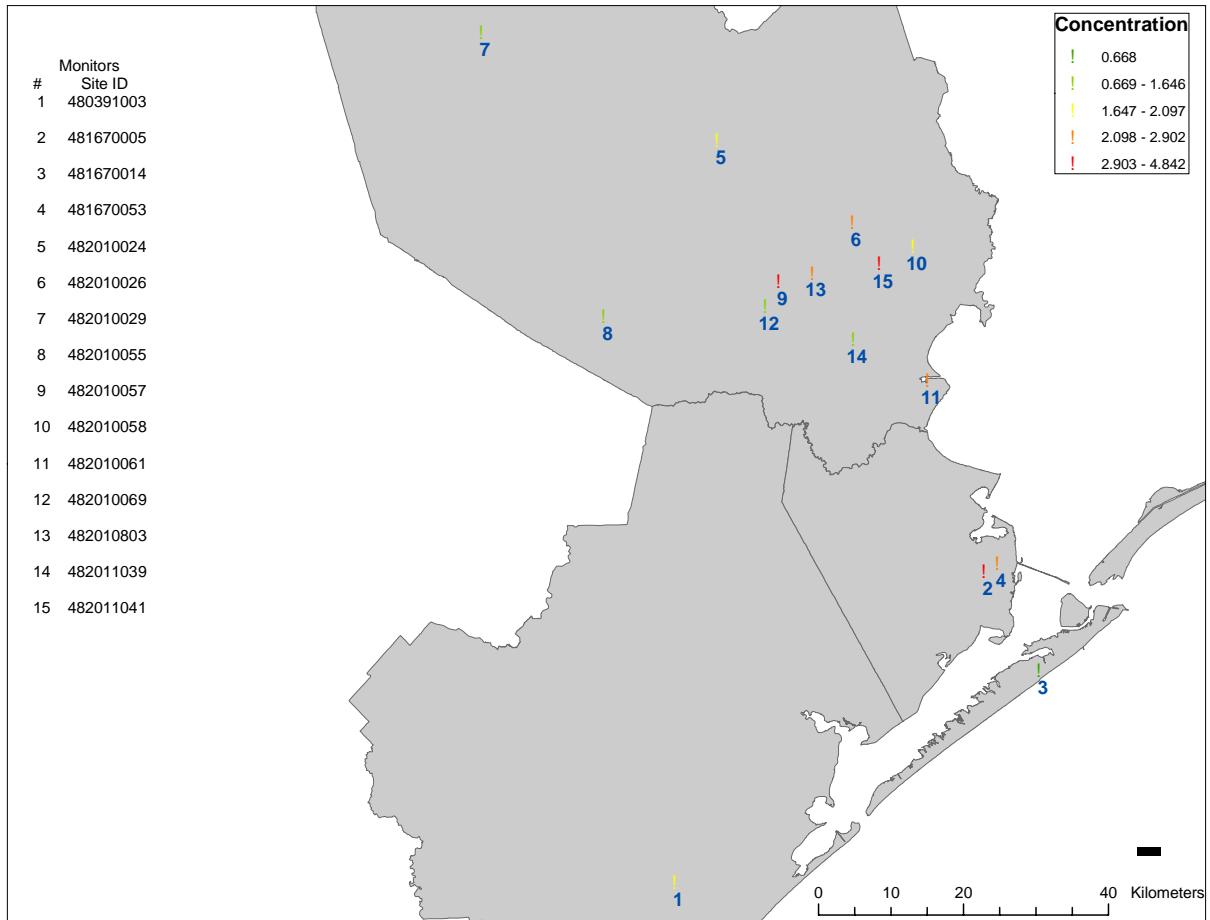
Figure 10. Monitor locations and annual average concentrations (μg per m^3) for the year 2000.
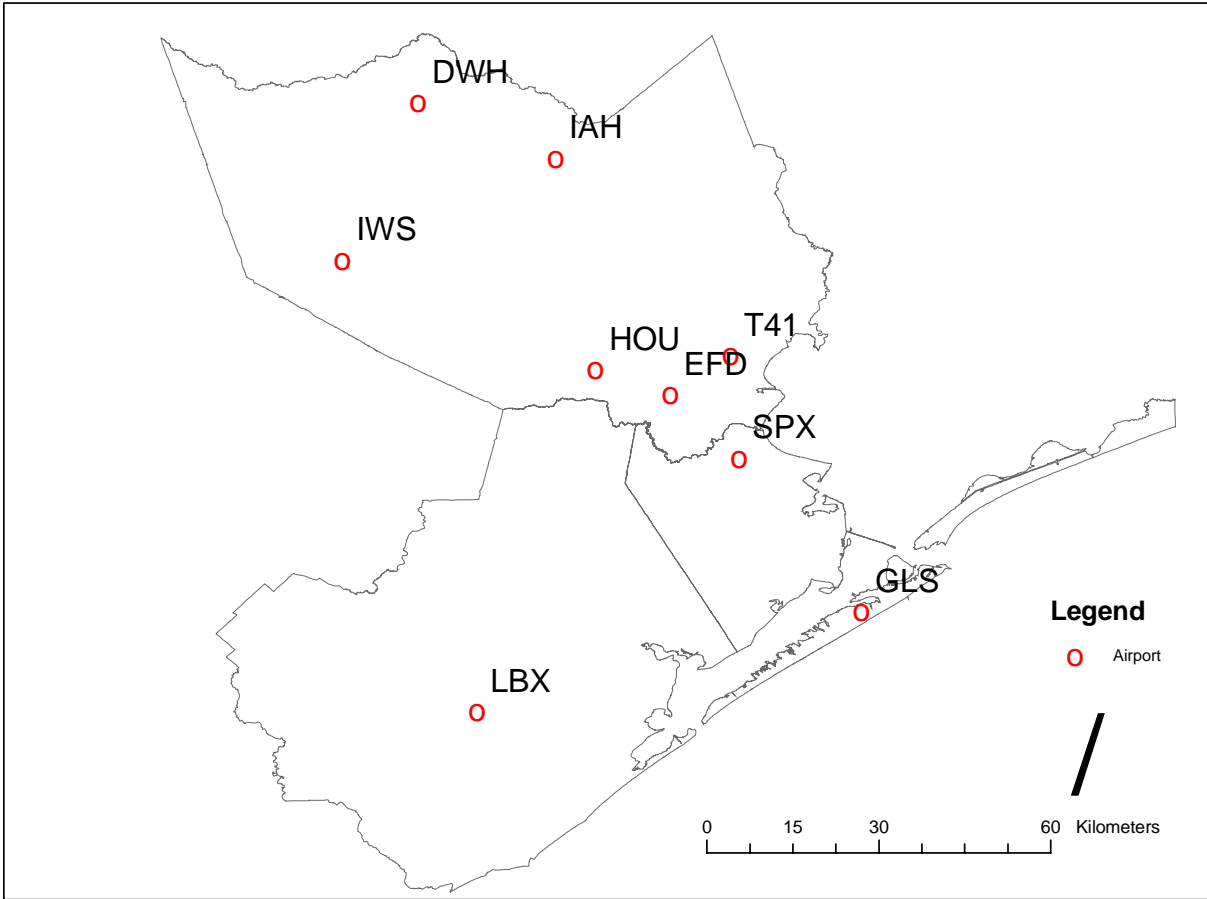
Figure 11. Locations of airports in Houston domain.

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
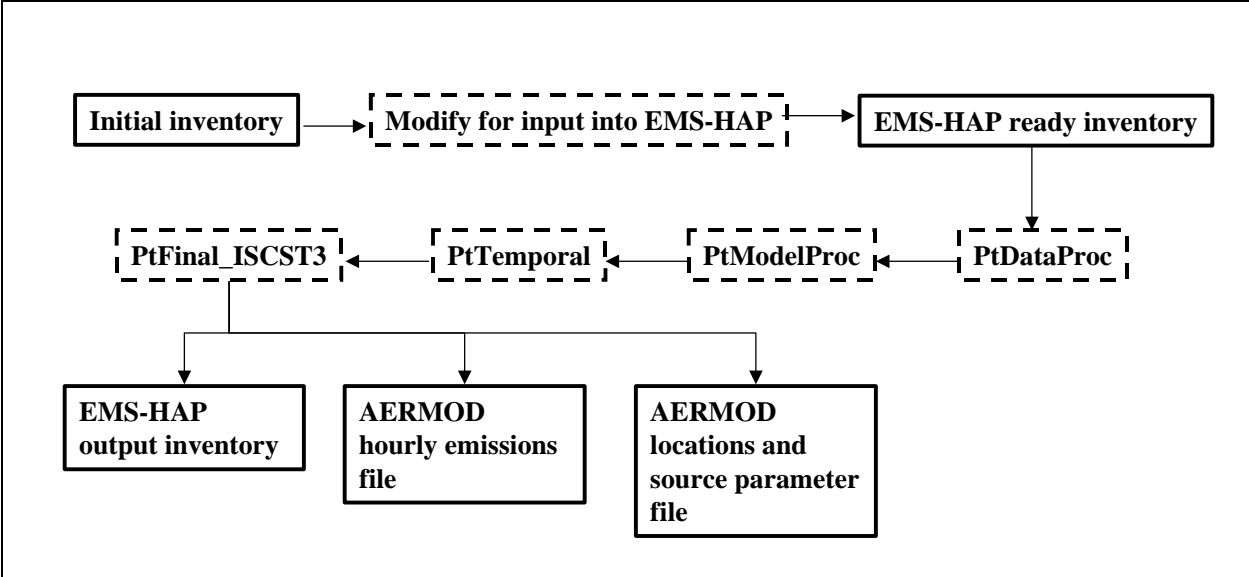
Figure 12. Point inventory EMS-HAP processing.
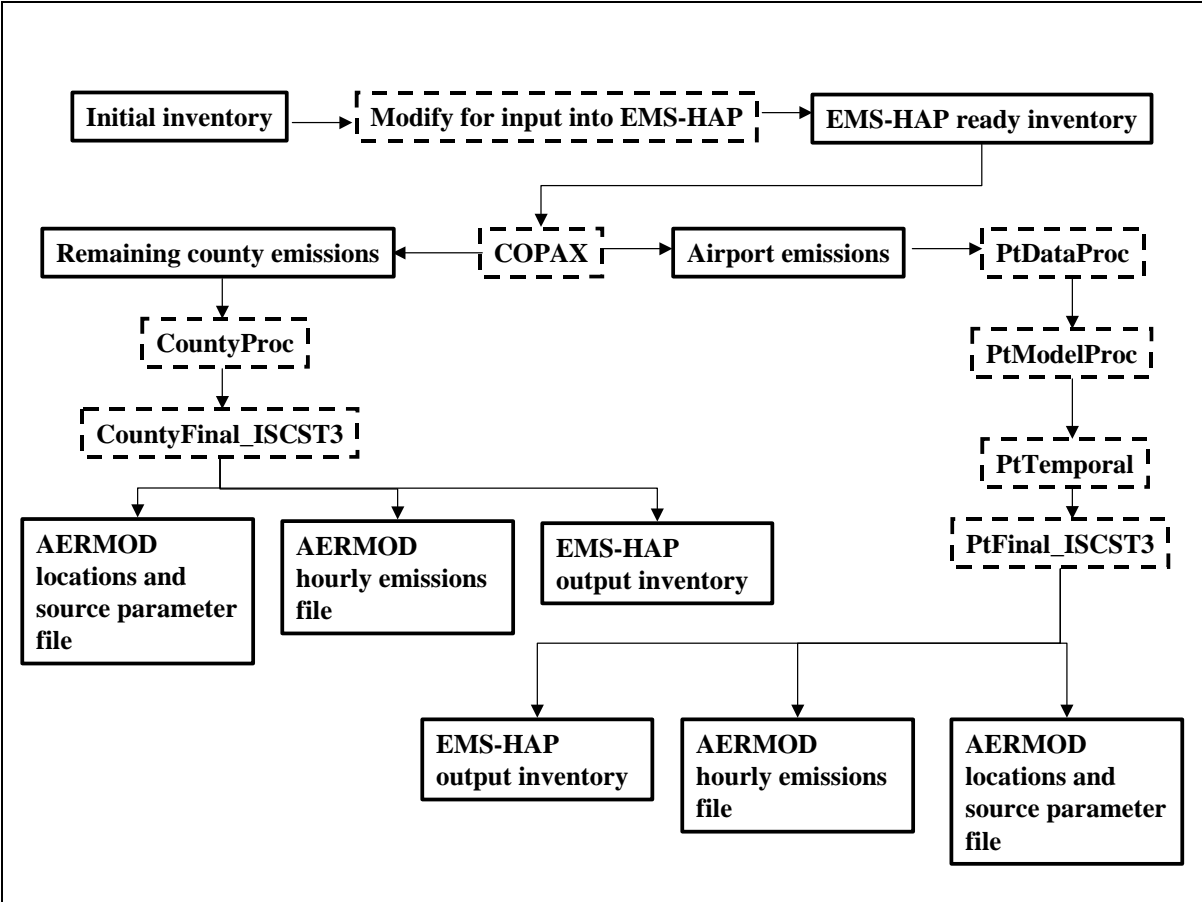
Figure 13. Emissions processing of nonpoint and nonroad inventories.
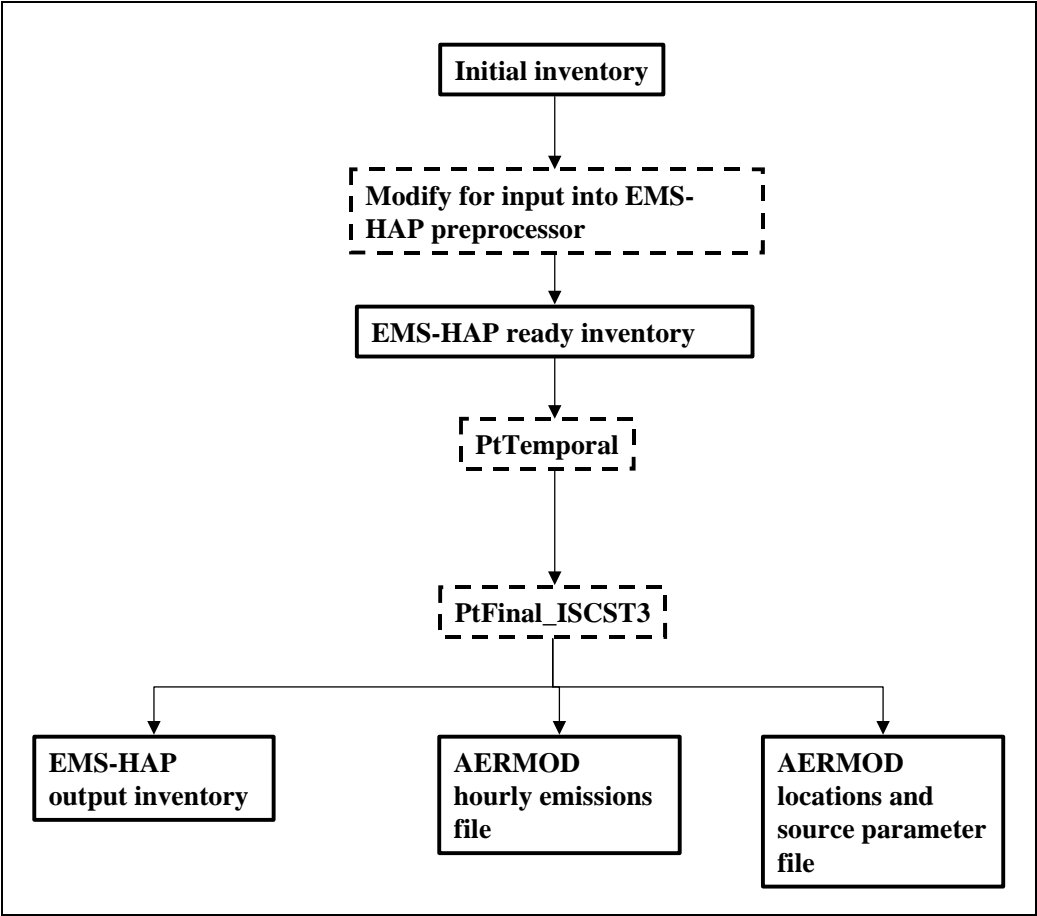
Figure 14. Onroad inventory emissions processing.
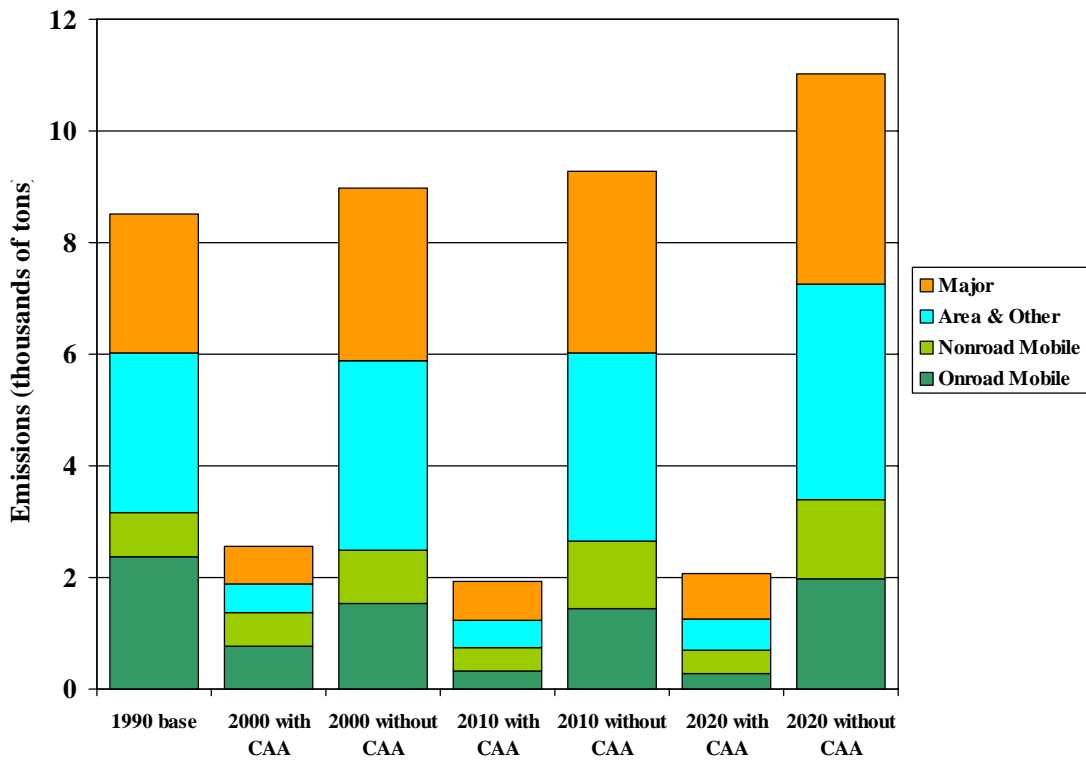
Figure 15. Major, area & other, onroad, and nonroad emissions (tons) for each year and inventory type.
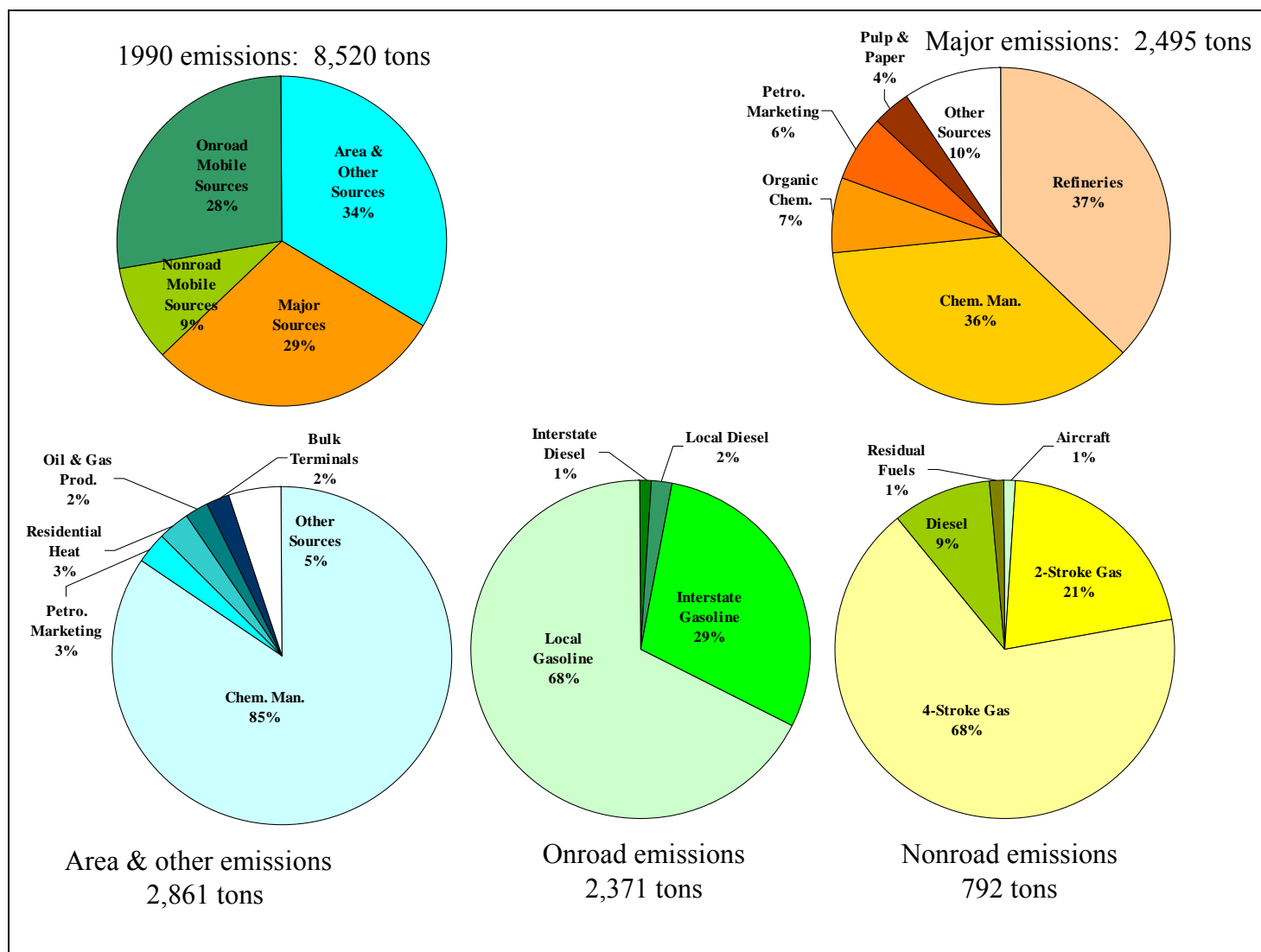
Figure 16. Breakdown of 1990 emissions by source category and within each source category. "Chemical Manufacturing" is included in the area source inventory. With-CAA emissions include Chemical Manufacturing in the point (major source) inventory.

2000 emissions: 2,557 tons

- Area & Other Sources 20%
- Major Sources 27%
- Nonroad Mobile Sources 24%
- Onroad Mobile Sources 29%

Major emissions: 681 tons

- Bulk Terminals 3%
- Other Sources 11%
- Organic Chem. 6%
- Chem. Man. 46%
- Refineries 14%
- Petro. Ind. 20%

Area & other emissions 499 tons

- Petro Marketing 24%
- Organic Chem. 10%
- Bulk Terminals 10%
- Fires 10%
- Residential Heat 8%
- Other Sources 38%

Onroad emissions 759 tons

- Interstate Diesel 1%
- Local Diesel 2%
- Interstate Gasoline 31%
- Local Gasoline 66%

Nonroad emissions 617 tons

- Residual Fuels 0%
- Aircraft 2%
- Diesel 13%
- 2-Stroke Gas 28%
- 4-Stroke Gas 57%

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.
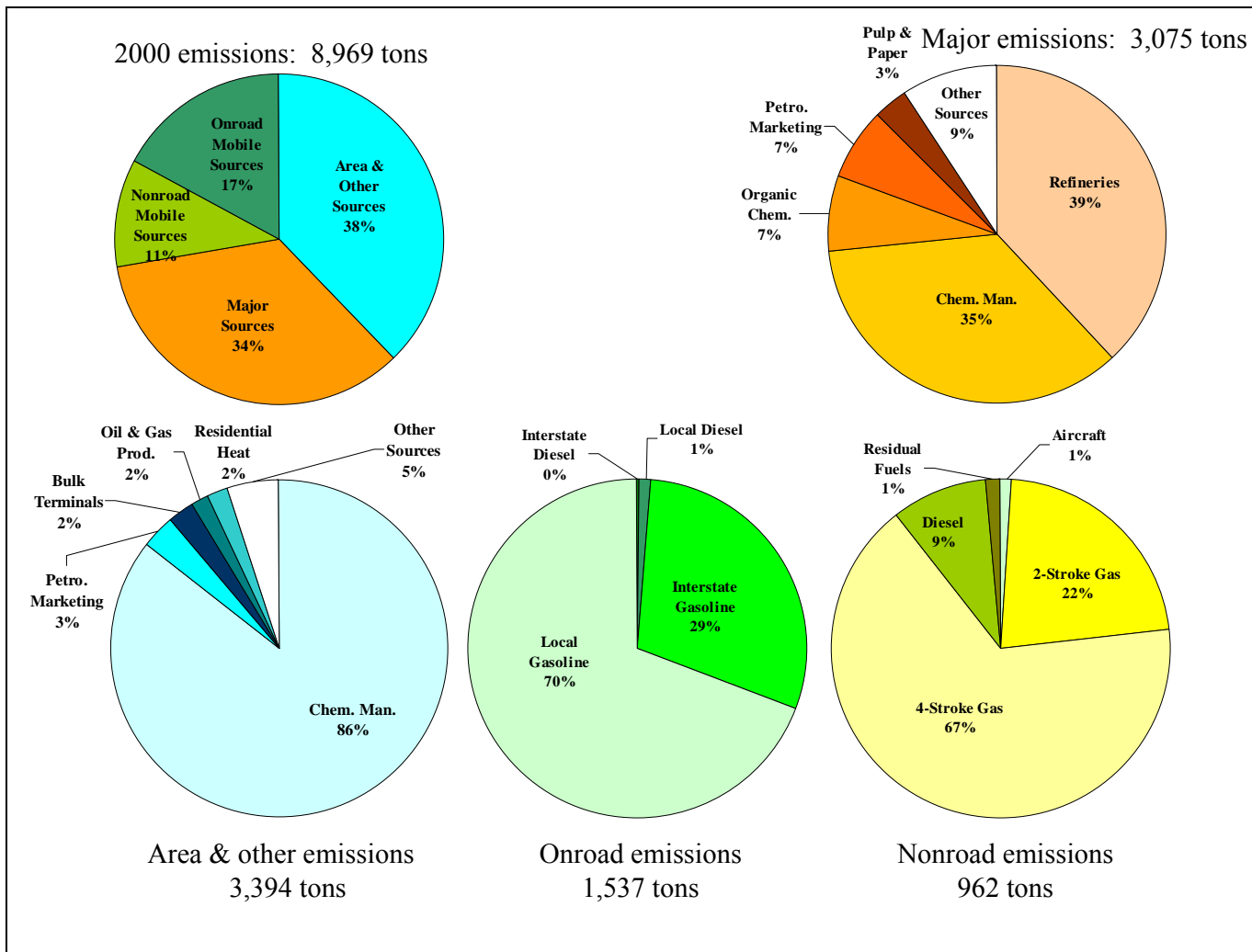
Figure 18. Breakdown of 2000 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.
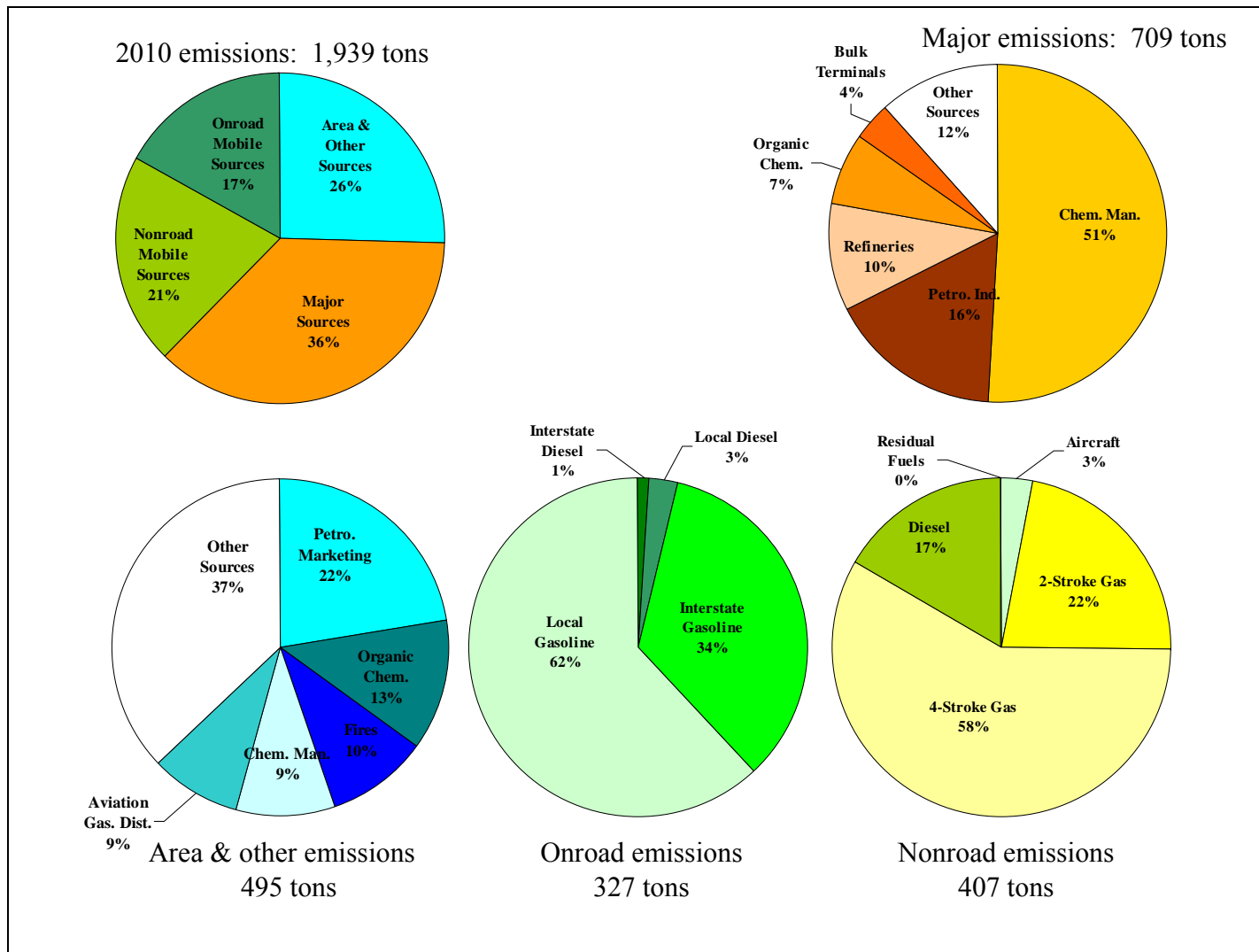
Figure 19. Breakdown of 2010 CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.
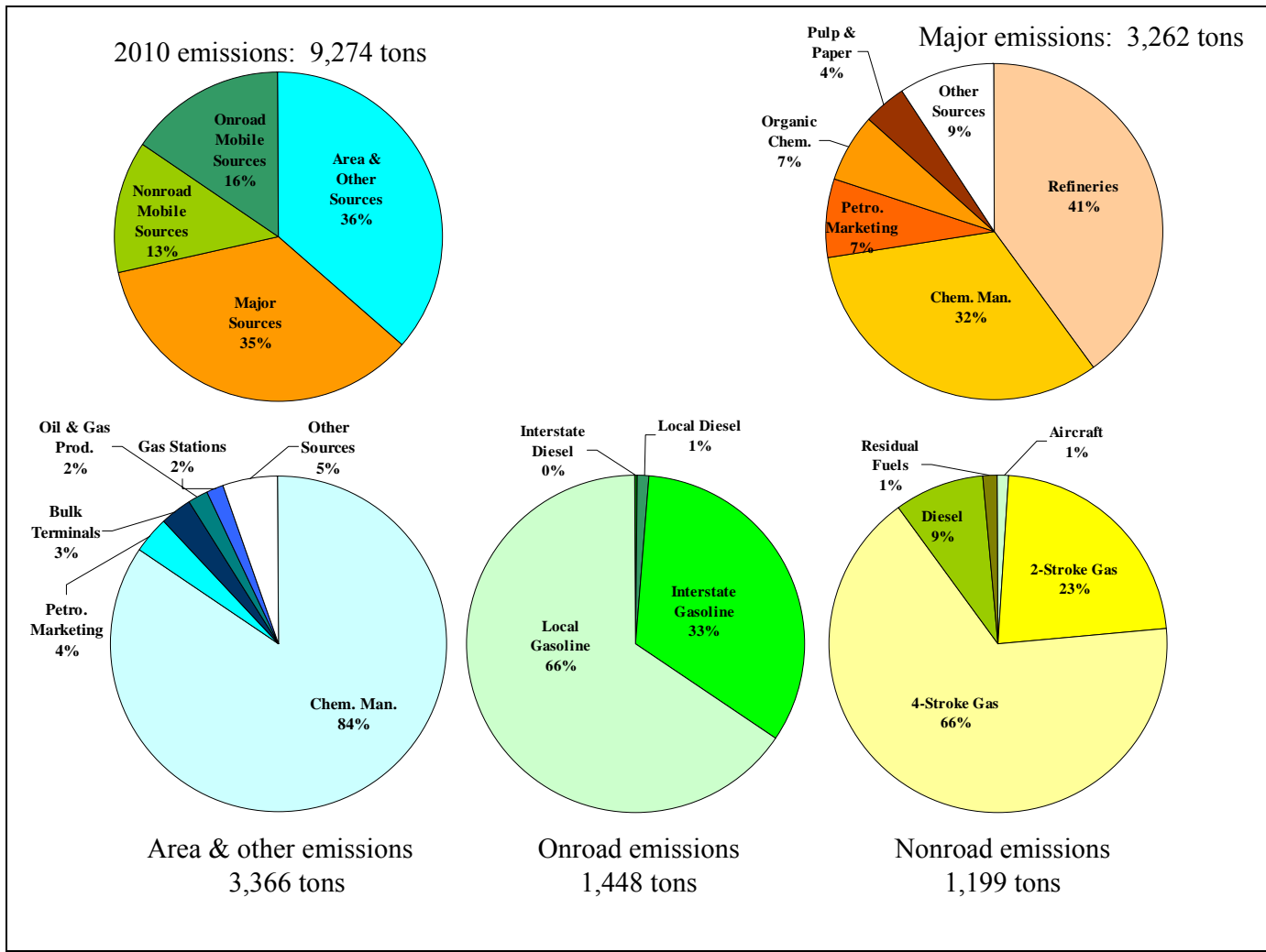
Figure 20. Breakdown of 2010 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

63
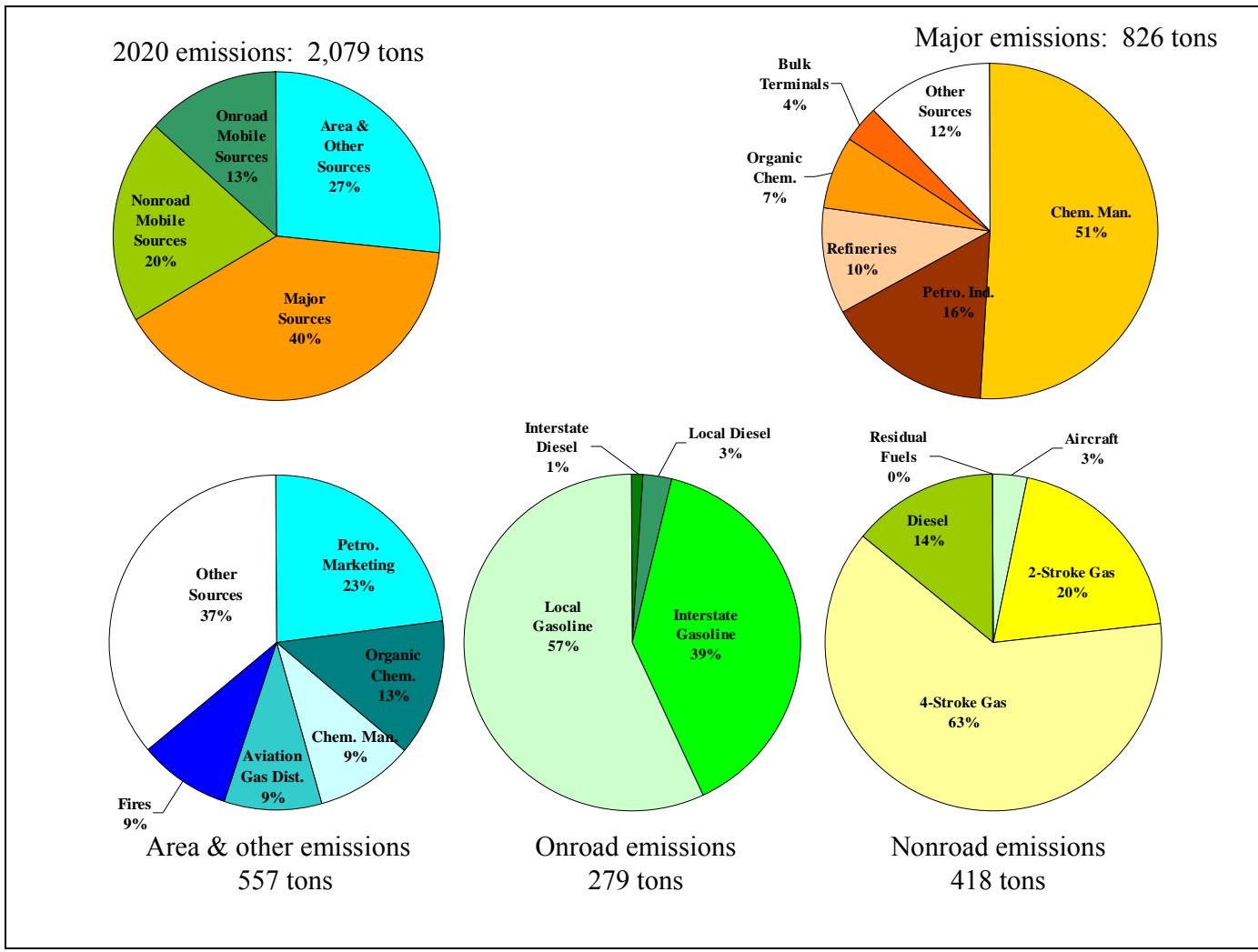
Figure 21. Breakdown of 2020 CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.

2020 emissions: 11,025 tons

- Area & Other Sources 35%
- Major Sources 34%
- Nonroad Mobile Sources 13%
- Onroad Mobile Sources 18%

Major emissions: 3,775 tons

- Refineries 39%
- Chem. Man. 32%
- Other Sources 9%
- Petro. Marketing 8%
- Organic Chem. 7%
- Pulp & Paper 5%

Area & other emissions 3,846 tons

- Chem. Man. 84%
- Other Sources 5%
- Petro. Marketing 4%
- Bulk Terminals 3%
- Gas Stations 2%
- Oil & Gas Prod. 2%

Onroad emissions 1,971 tons

- Local Gasoline 61%
- Interstate Gasoline 38%
- Local Diesel 1%
- Interstate Diesel 0%

Nonroad emissions 1,433 tons

- 4-Stroke Gas 68%
- 2-Stroke Gas 22%
- Diesel 8%
- Residual Fuels 1%
- Aircraft 1%
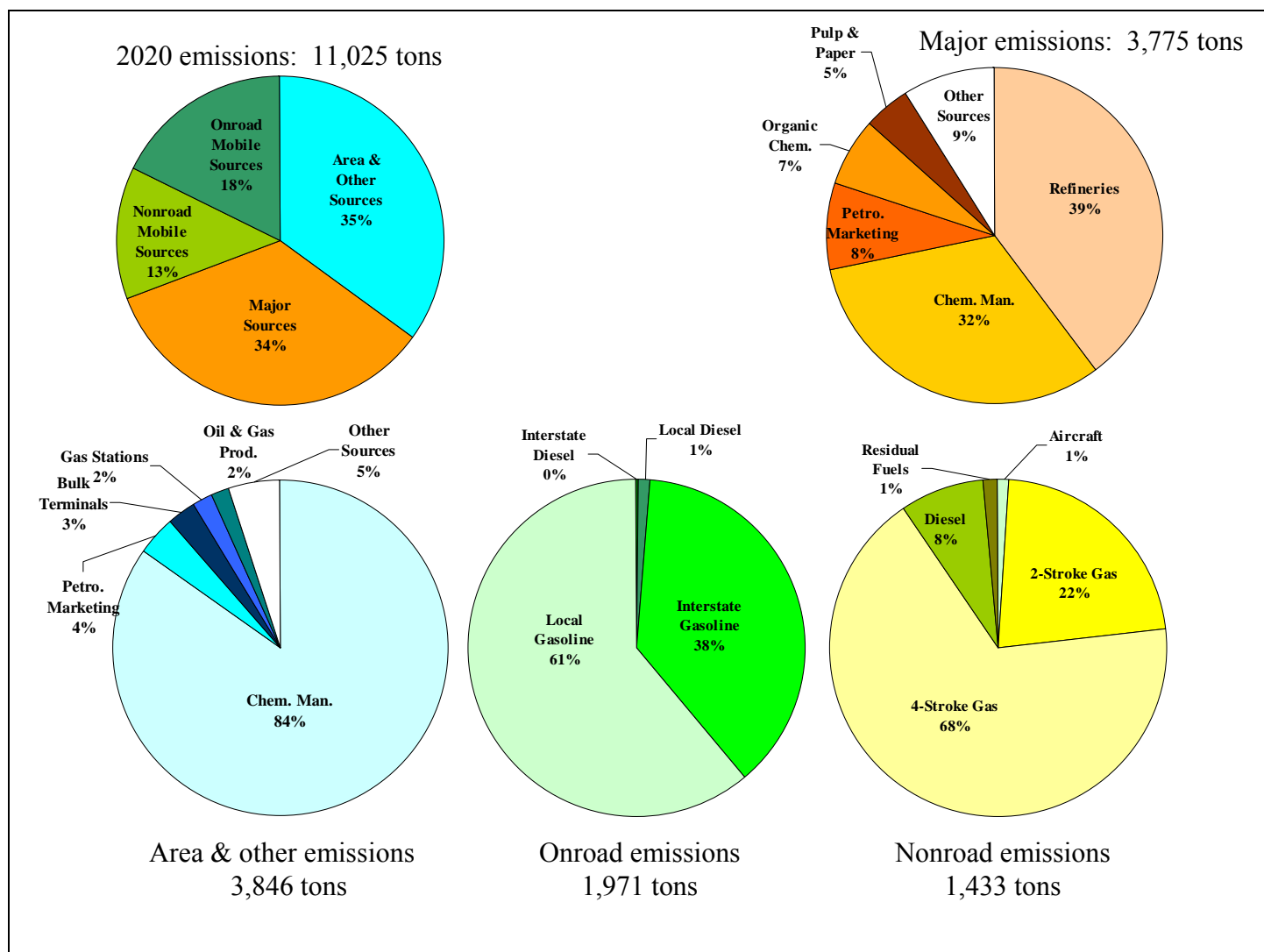
Figure 22. Breakdown of 2020 NON-CAA emissions by source category and within each source category. As noted in section 4.5 above, all NON-CAA emissions include "Chemical Manufacturing" in the area source (nonpoint) inventory, while the CAA emissions include Chemical Manufacturing in the major source (point) inventory.
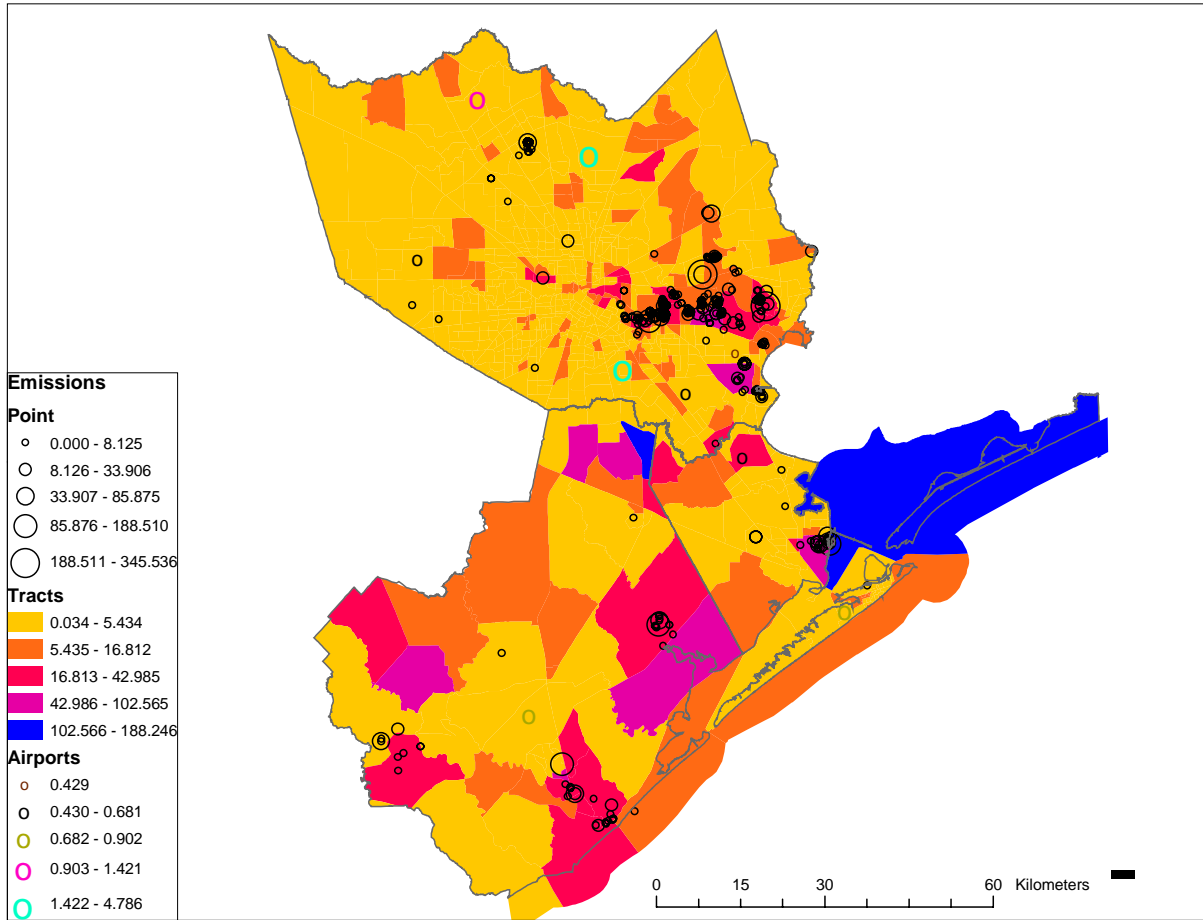
Figure 23. 1990 point (circles), airports (airplane symbols), and spatially allocated total tract (nonpoint plus nonroad) emissions. Emissions are in tons.

66
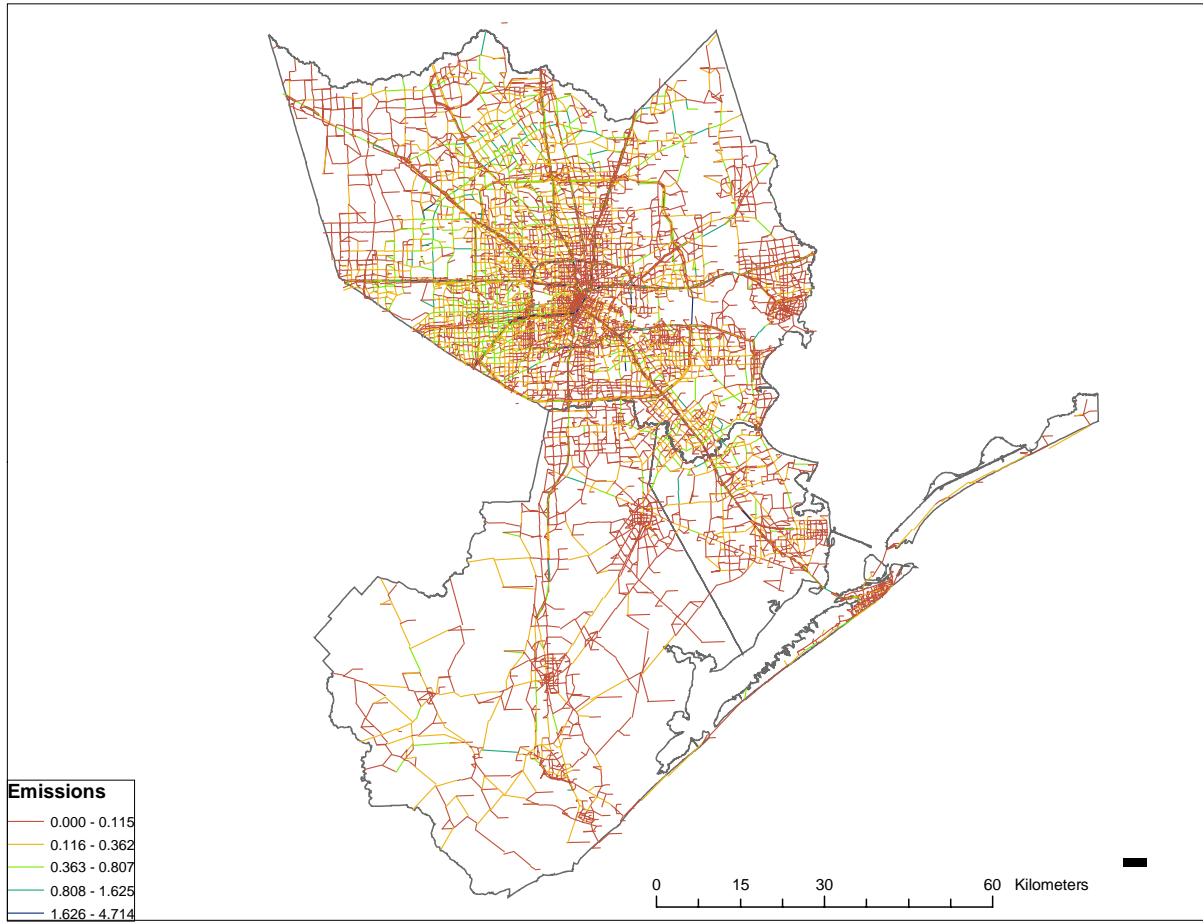
Figure 24. 1990 onroad emissions in tons.
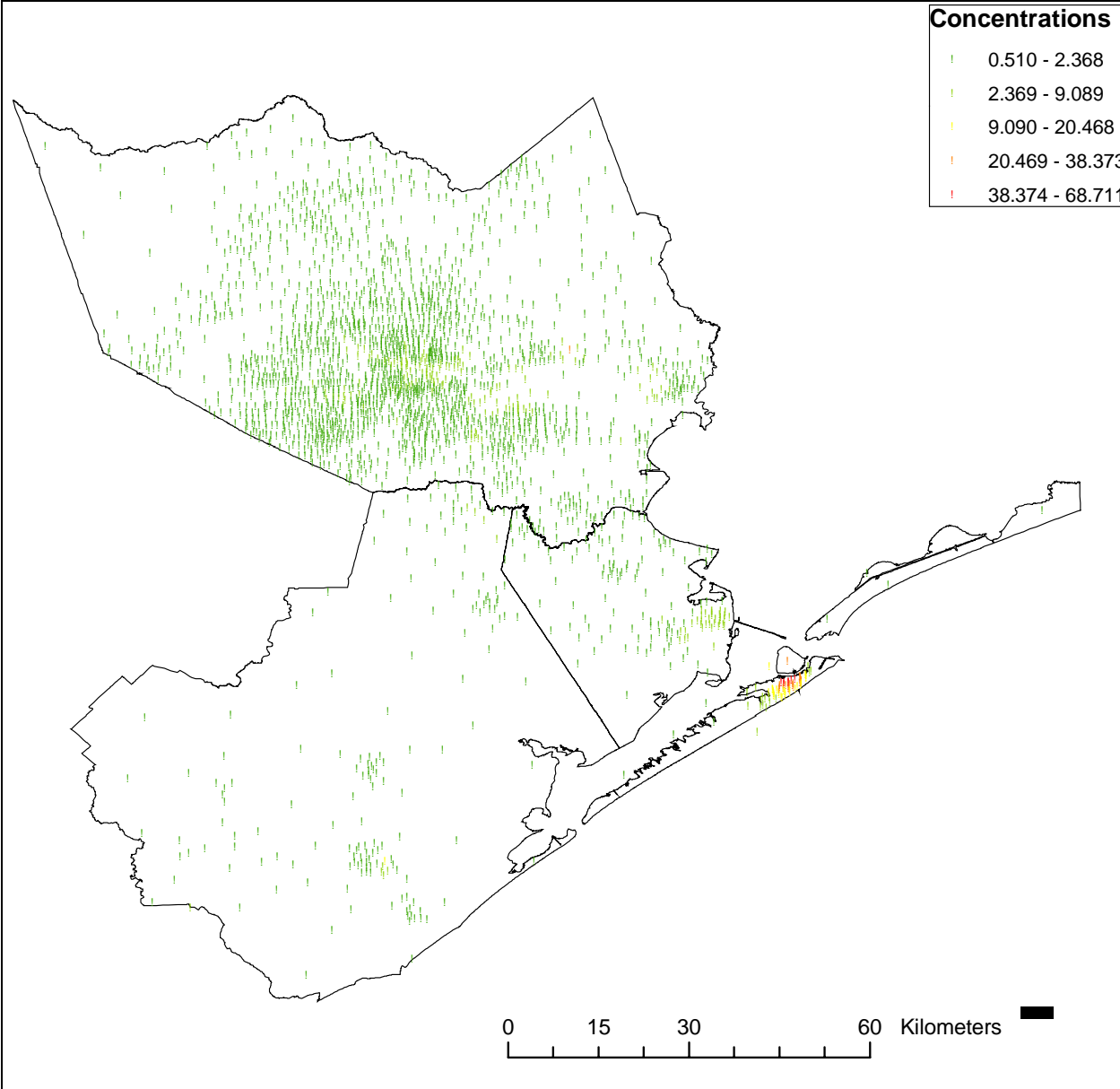
Figure 25. Annual average AERMOD concentrations (µg per m^3) at the census tract block level (with background included) for year 1990.
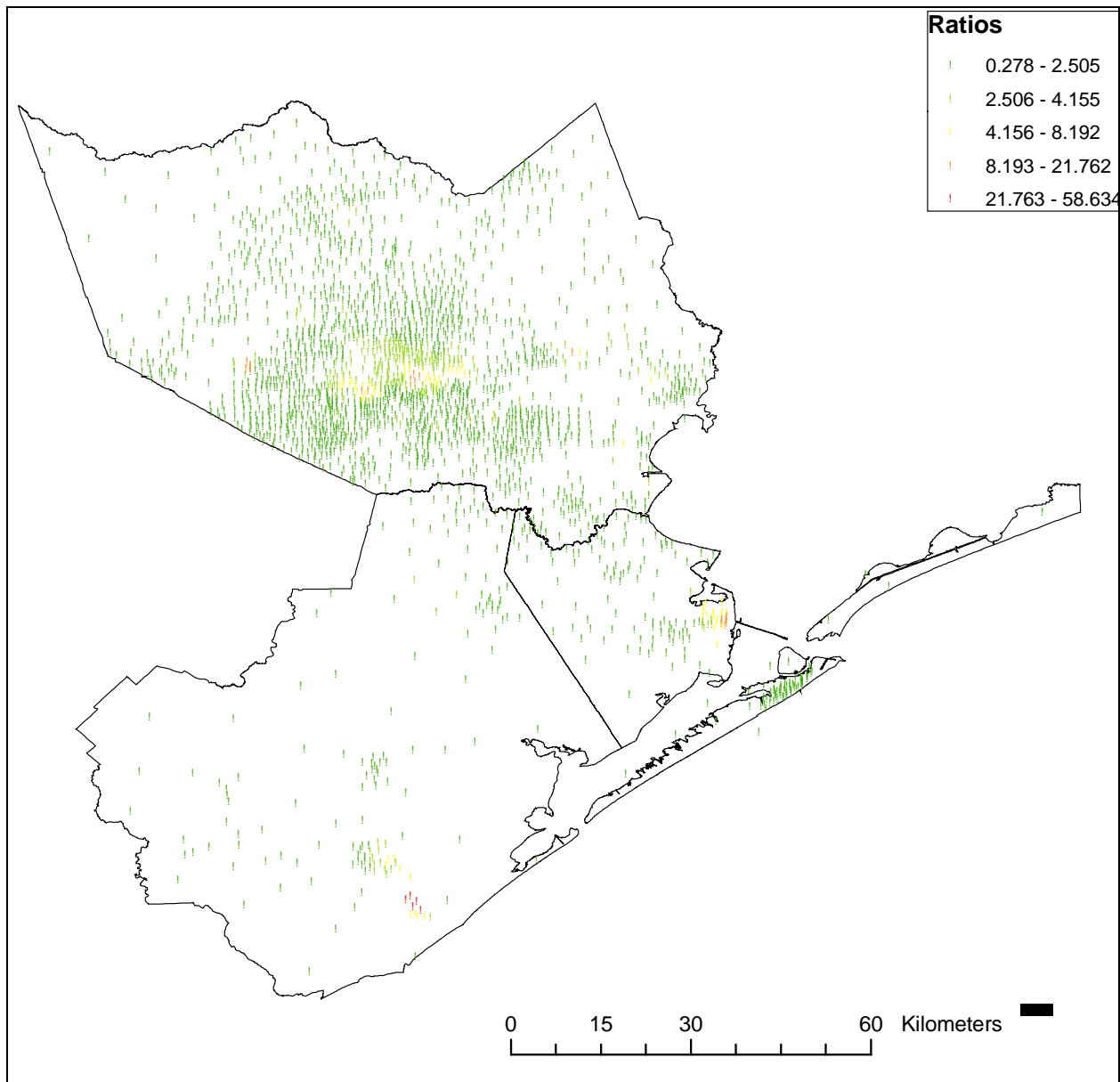
Figure 26. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2000. Concentrations are total concentrations with background included.
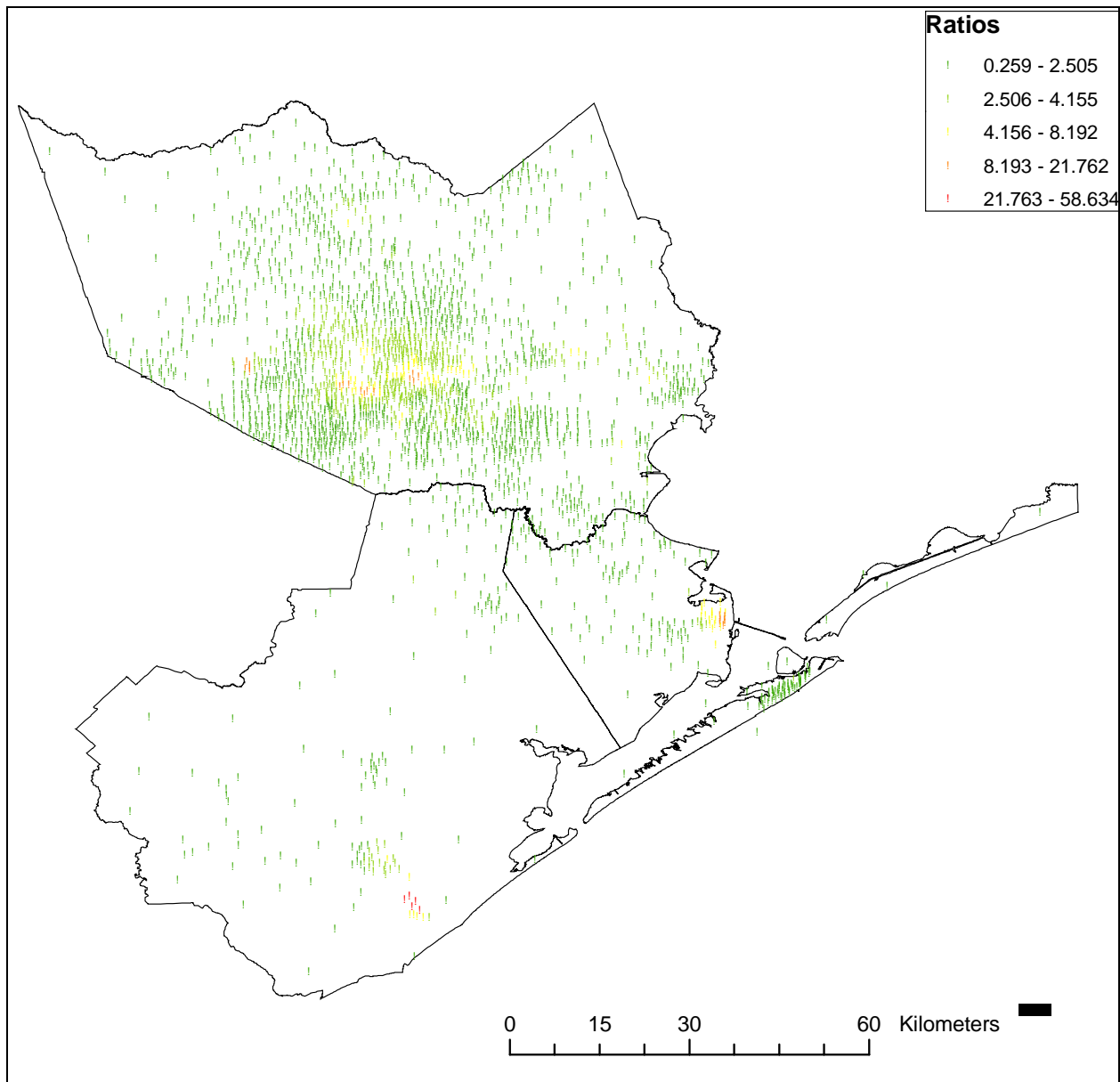
Figure 27. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2010. Concentrations are total concentrations with background included.

Ratios

	0.254 - 2.505
	2.506 - 4.155
	4.156 - 8.192
	8.193 - 21.762
	21.763 - 58.634

0   15   30   60 Kilometers
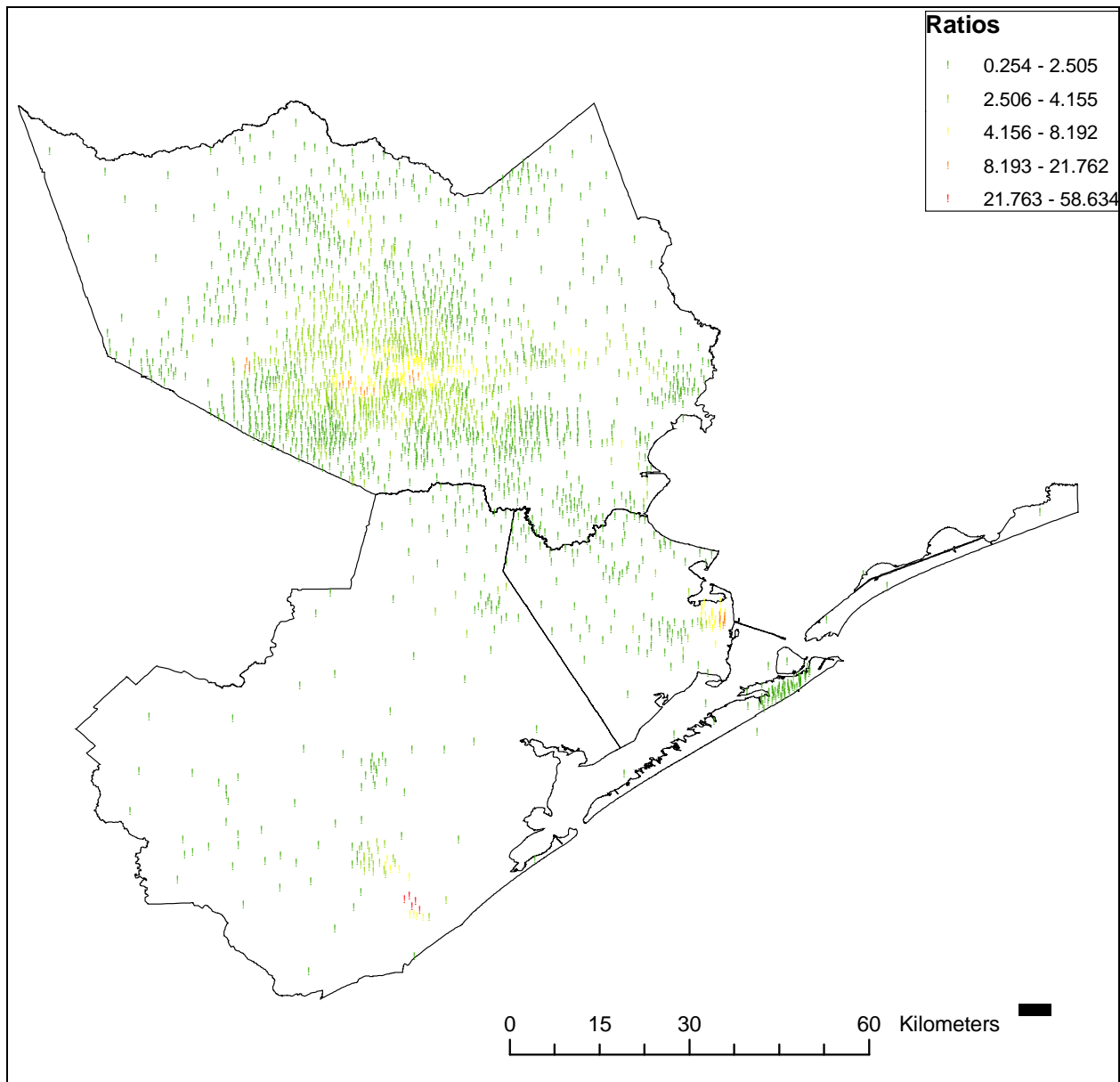
Figure 28. Ratio of AERMOD annual average NON-CAA concentrations to CAA concentrations at the census block group level for 2020. Concentrations are total concentrations with background included.
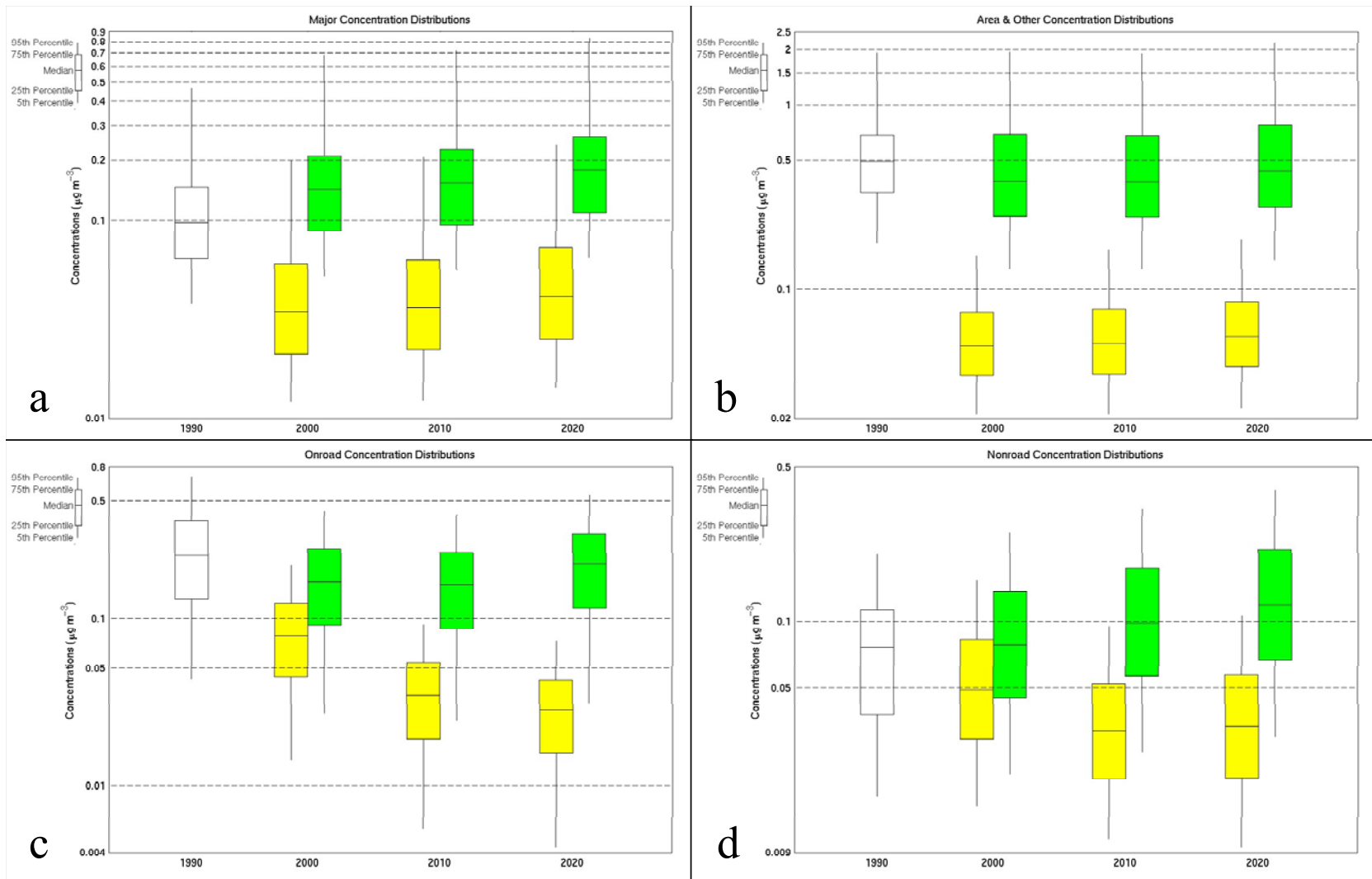
Figure 29. Box and whisker plots of block group level concentration (μg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. The white box denotes 1990, yellow boxes denote the CAA results and green boxes denote the NON-CAA results.
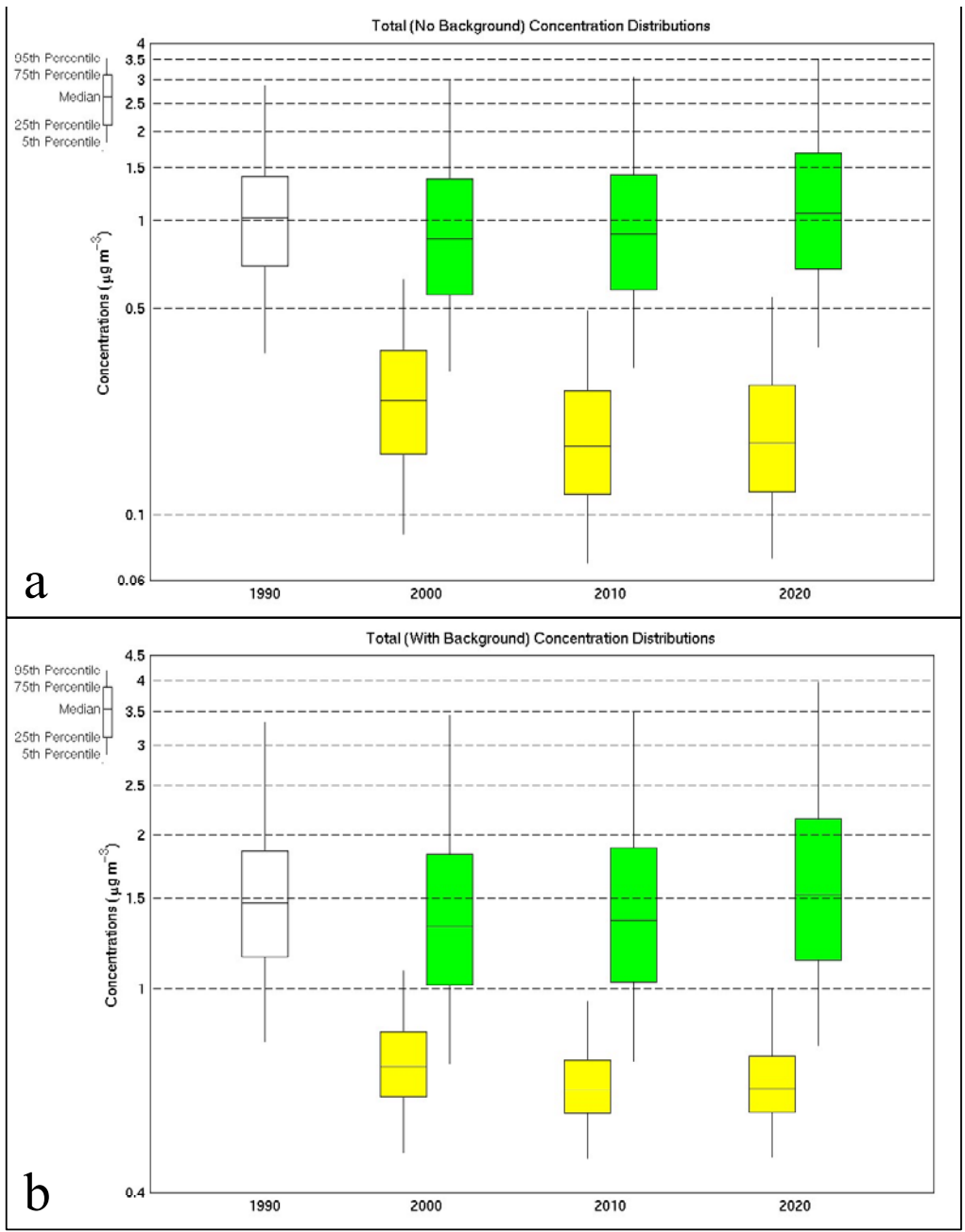
Figure 30. Box and whisker plots of block group level total concentration (µg per m³) distributions for 1990, 2000, 2010, and 2020 for a) without background and b) with background. Box colors are the same as in Figure 29.
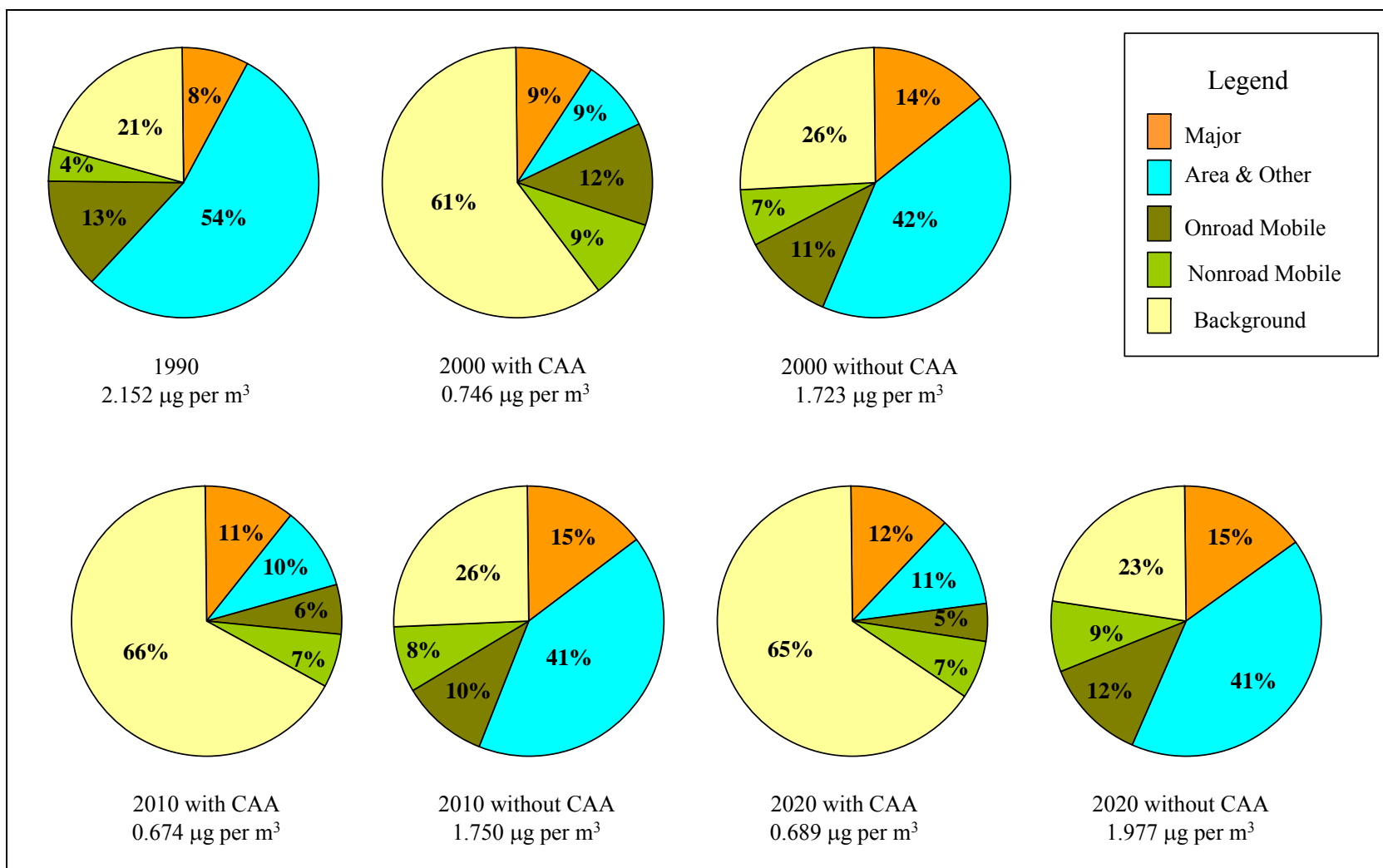
Figure 31. Percent contributions of each source category to the annual average AERMOD total concentrations averaged over all receptors. Total concentrations are given below each chart.

74
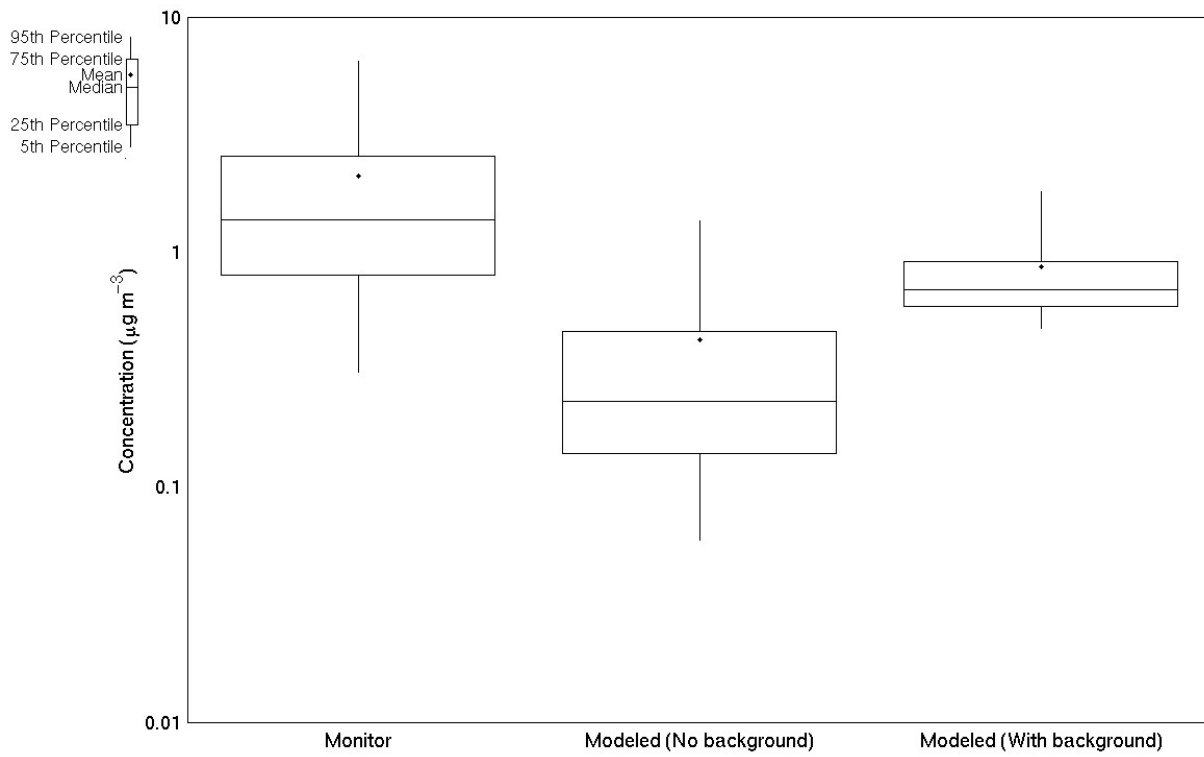
Figure 32. Box and whisker plots of daily monitor, modeled (without background), and modeled (with background) concentrations (μg per m^3). For the AERMOD model, only days with monitor values are included in the calculations.
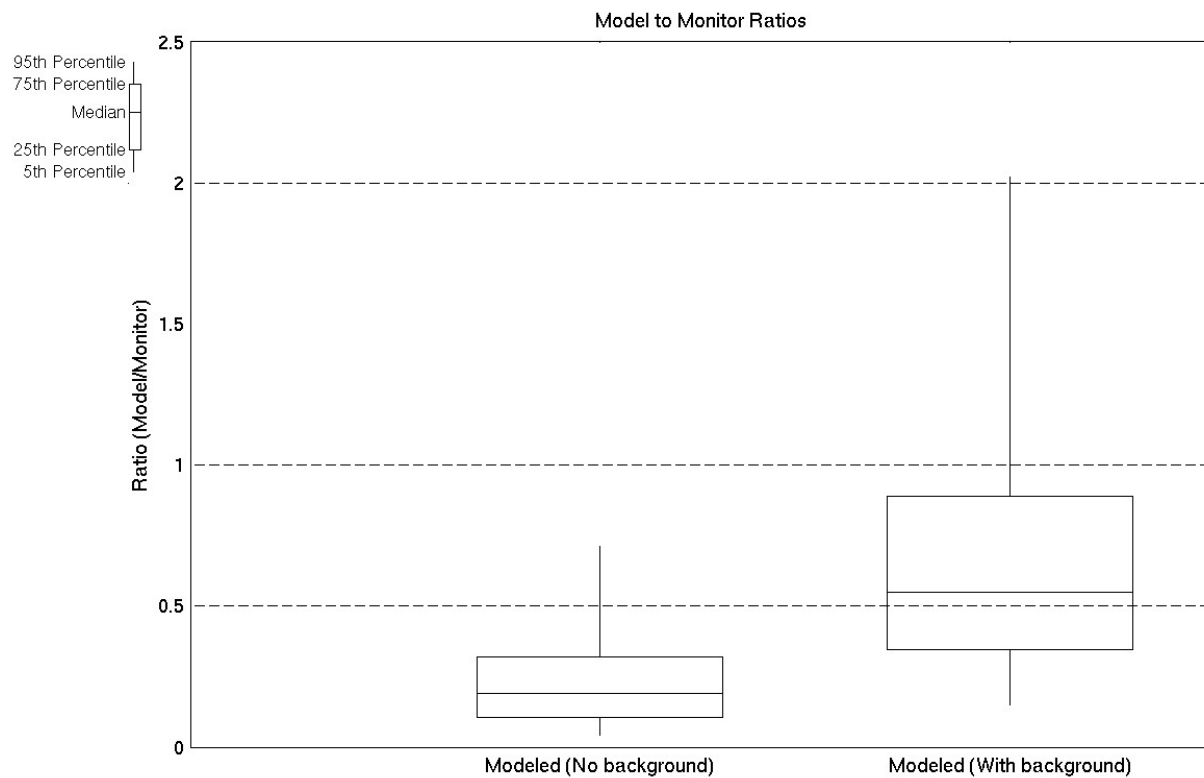
Figure 33. Ratios of daily 2000 CAA AERMOD modeled (with and without background) concentrations to observed monitor concentrations. For the AERMOD model, only days of monitor observations are included in the calculations.
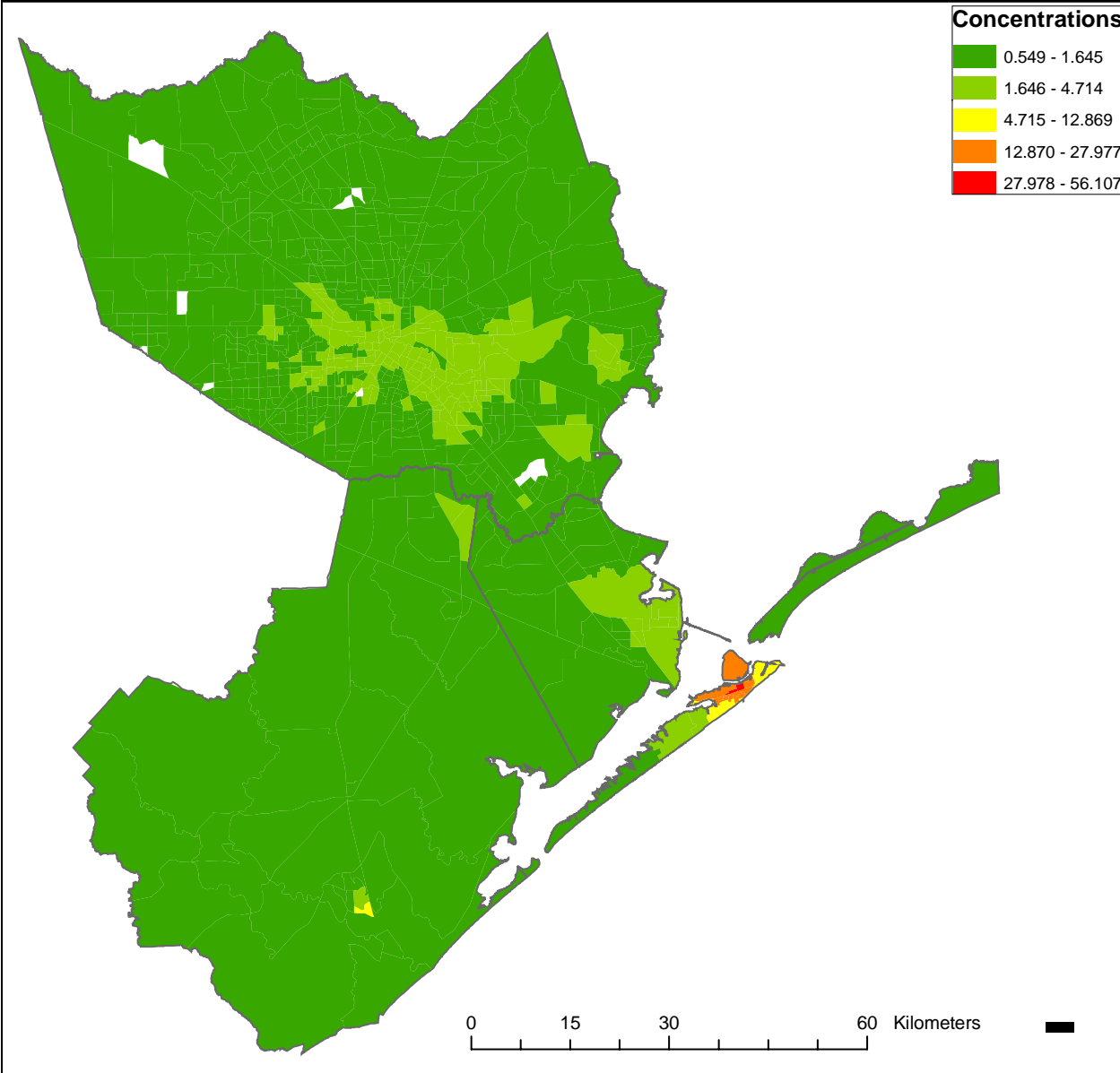
Figure 34. Annual average HAPEM6 concentrations (μg per m^3) at the census tract level (with background included) for year 1990.
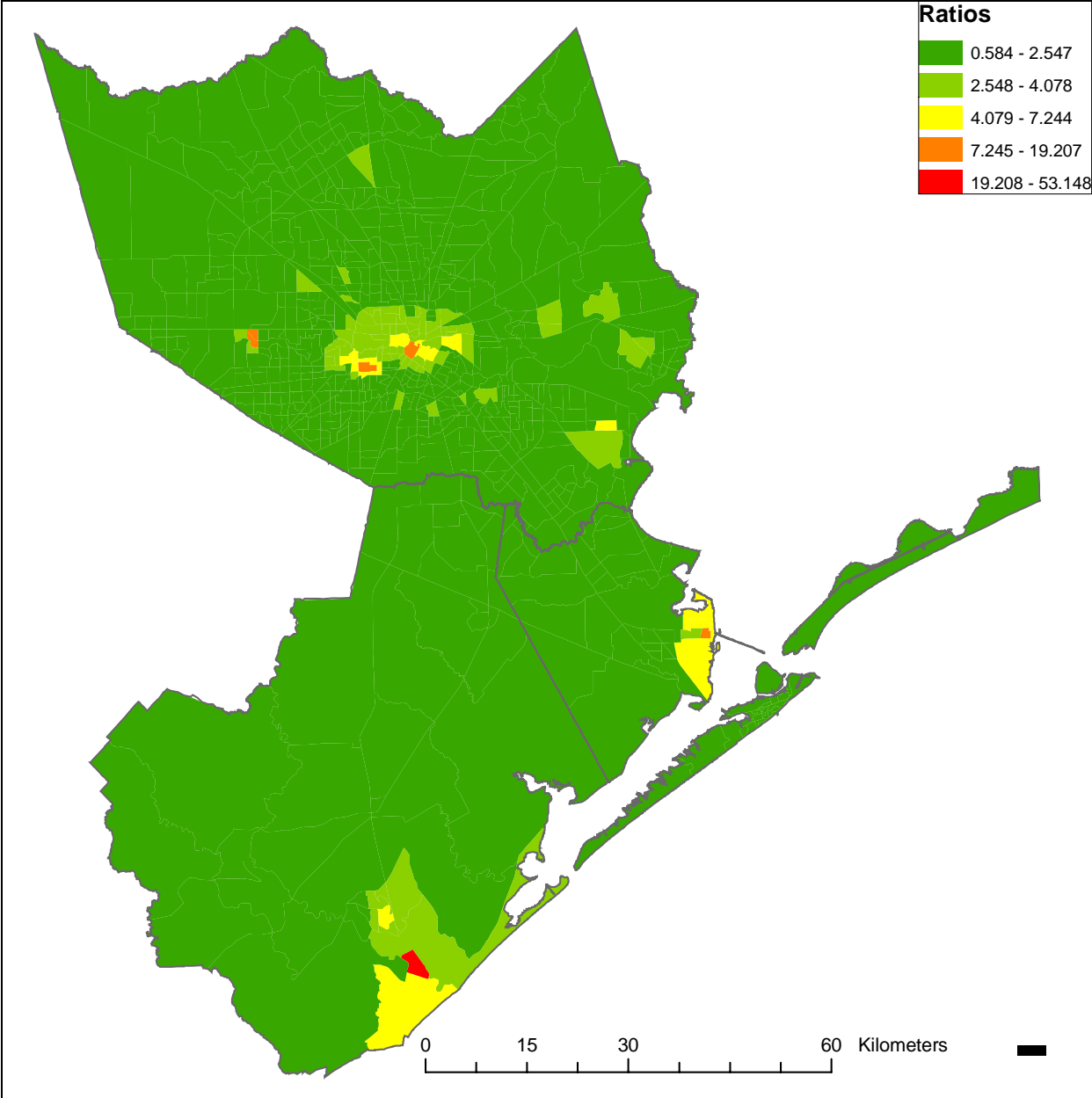
Figure 35. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2000. Concentrations are total concentrations with background included.
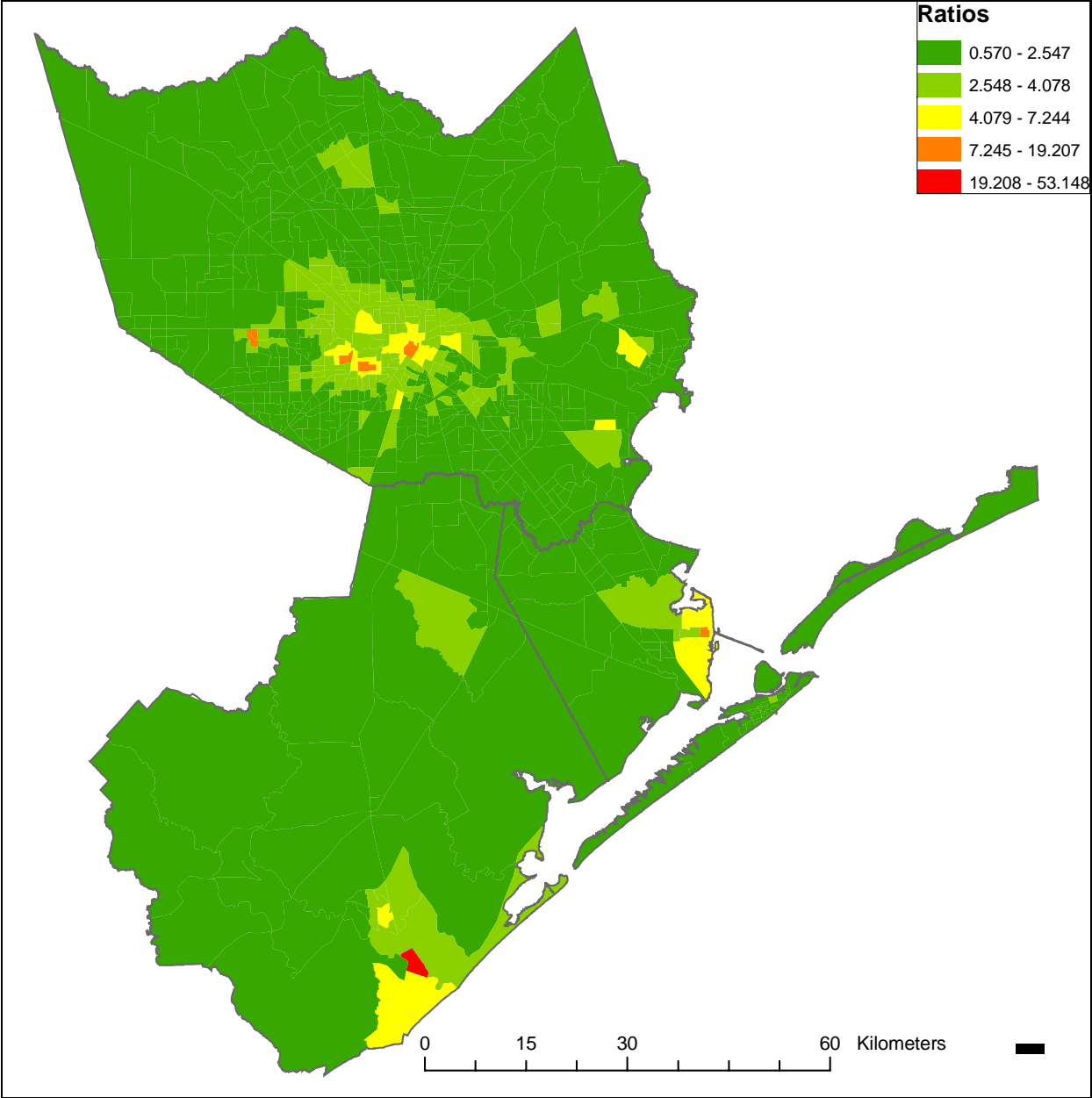
Figure 36. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2010. Concentrations are total concentrations with background included.
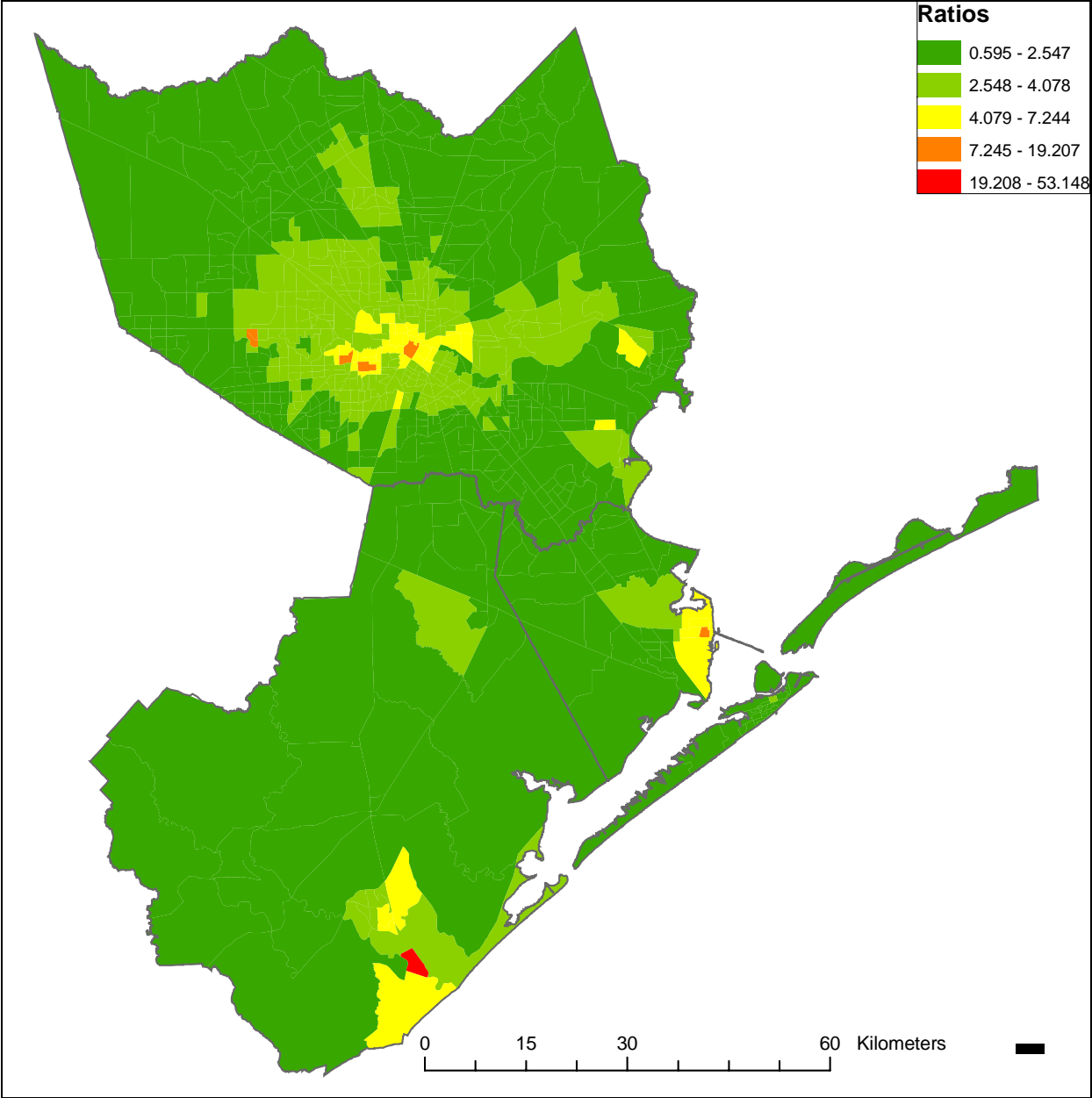
Figure 37. Ratio of HAPEM6 annual average NON-CAA concentrations to CAA concentrations at the census tract level for 2020. Concentrations are total concentrations with background included.
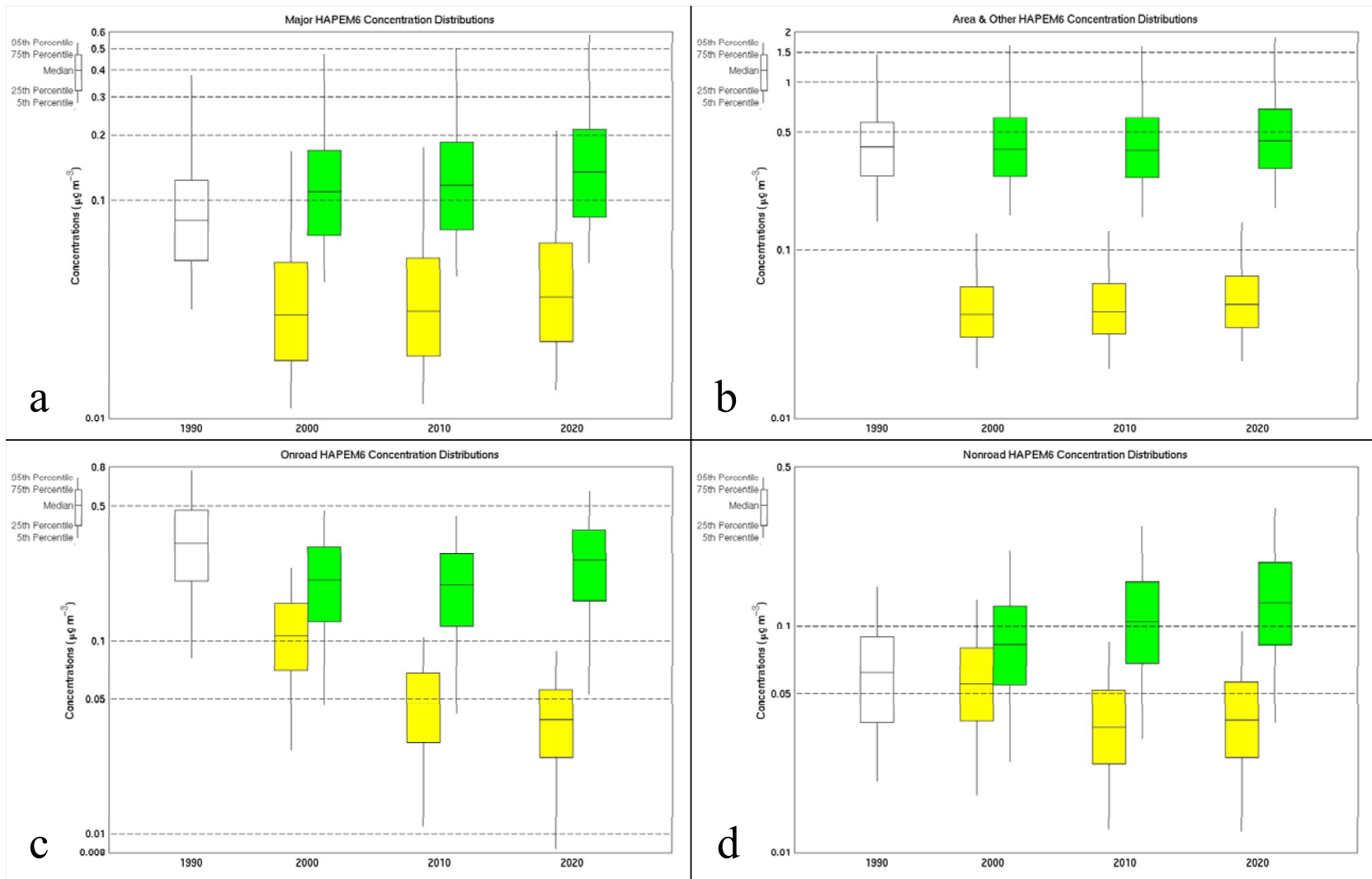
Figure 38. Box and whisker plots of tract level HAPEM6 concentration (μg per m³) distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. Box colors are the same as in Figure 29.
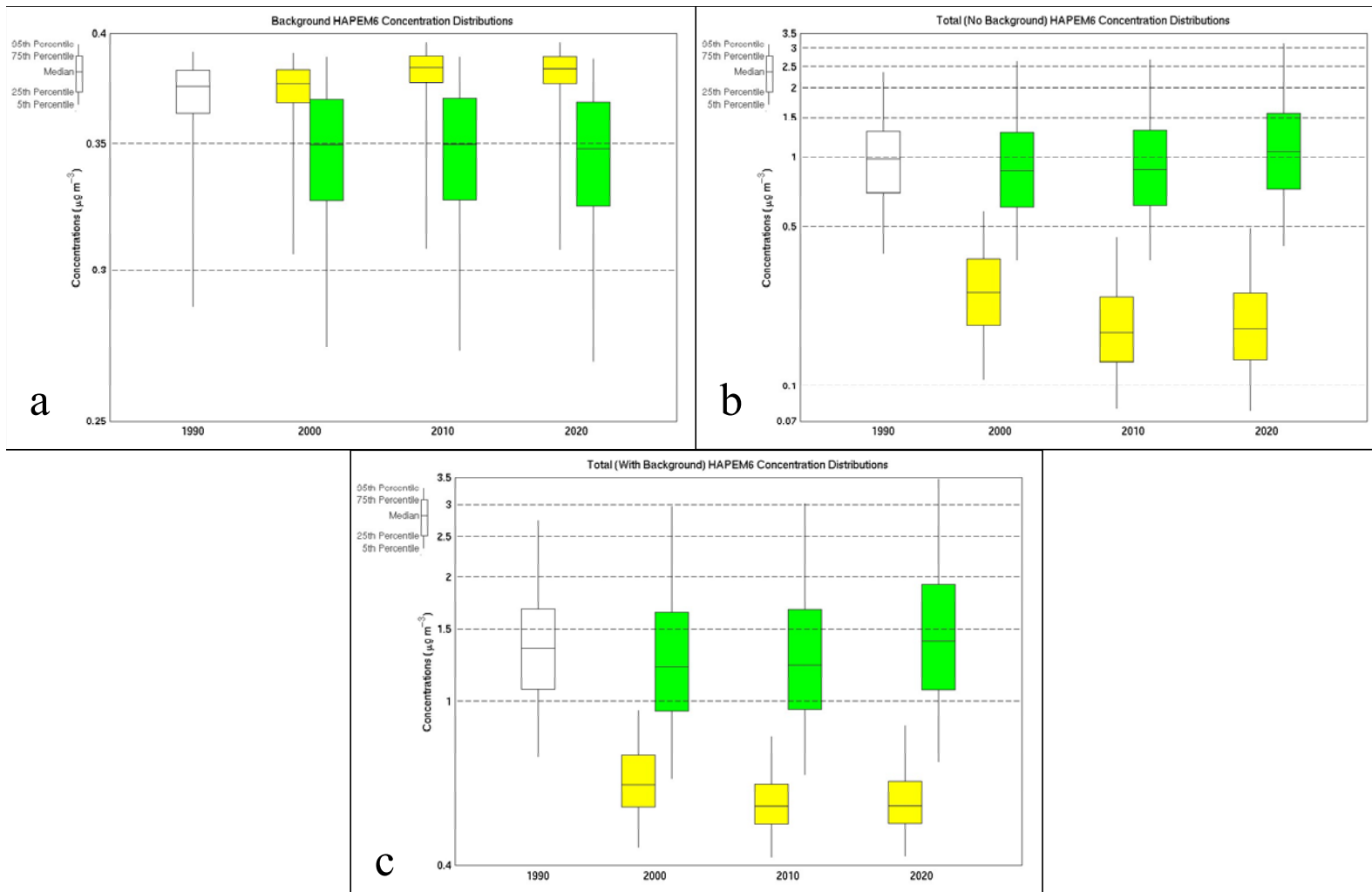
Figure 39. Box and whisker plots of tract level HAPEM6 concentration (μg per m^3) distributions for 1990, 2000, 2010, and 2020 for a) background, b) all sources not including background, and c) all sources including background included. Box colors are the same as in Figure 29.
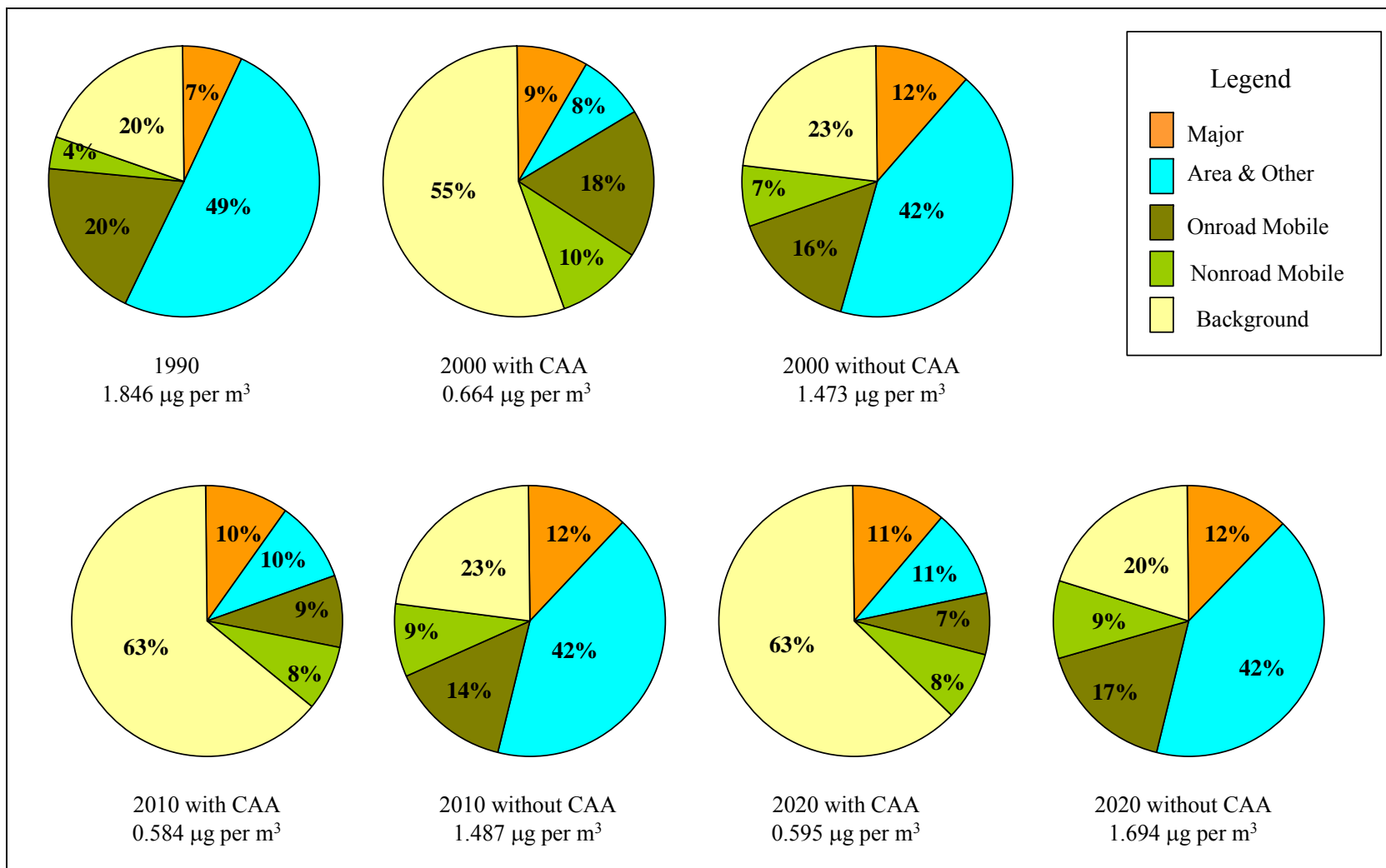
Figure 40. Percent contributions of each source category to the annual average HAPEM6 total concentrations averaged over all receptors. Total concentrations are given below each chart.

83
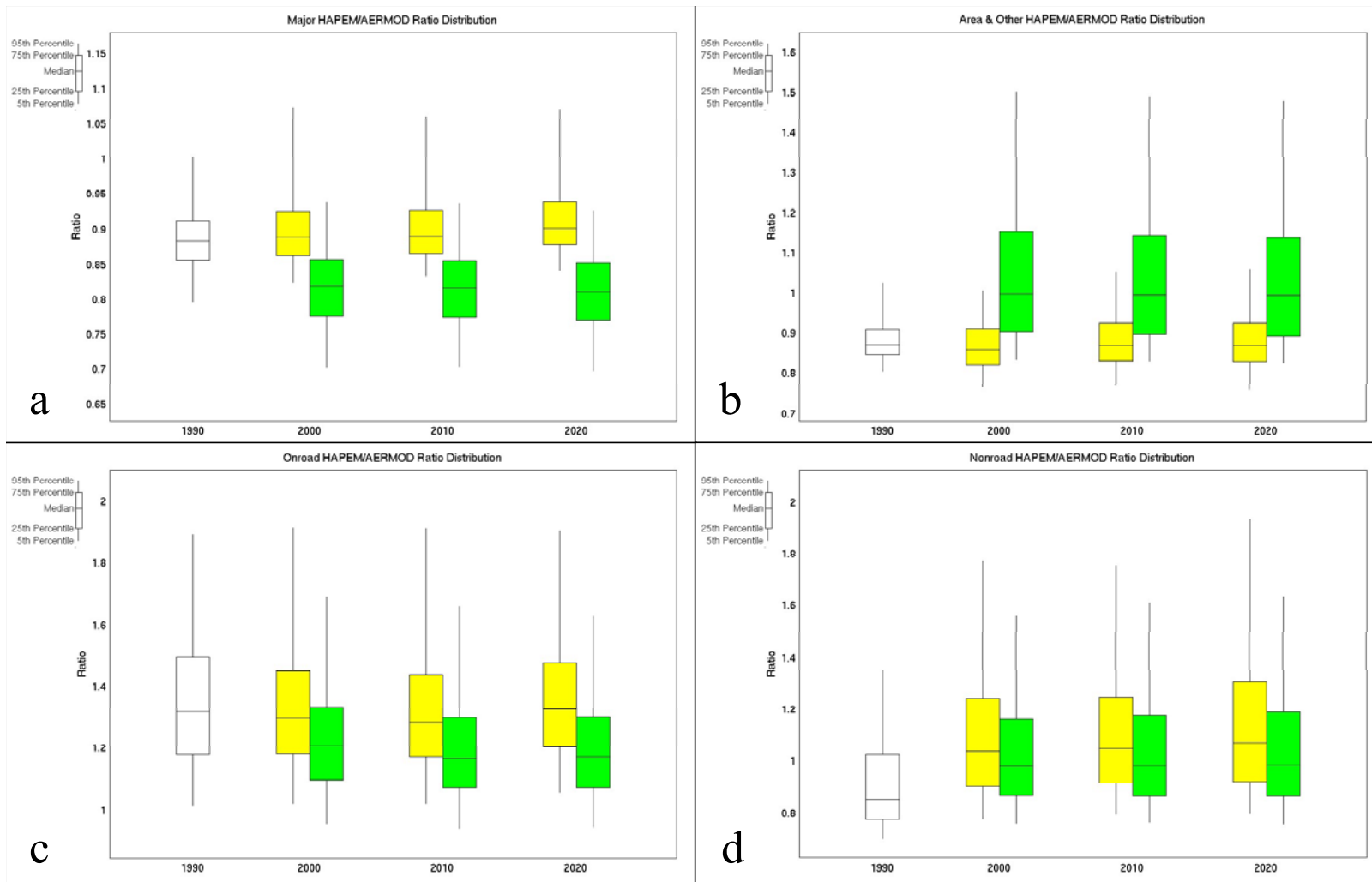
Figure 41. Box and whisker plots of tract level HAPEM6/AERMOD ratio distributions for 1990, 2000, 2010, and 2020 for a) major sources, b) area & other sources, c) onroad sources, and d) nonroad sources. Box colors are the same as in Figure 29.

84
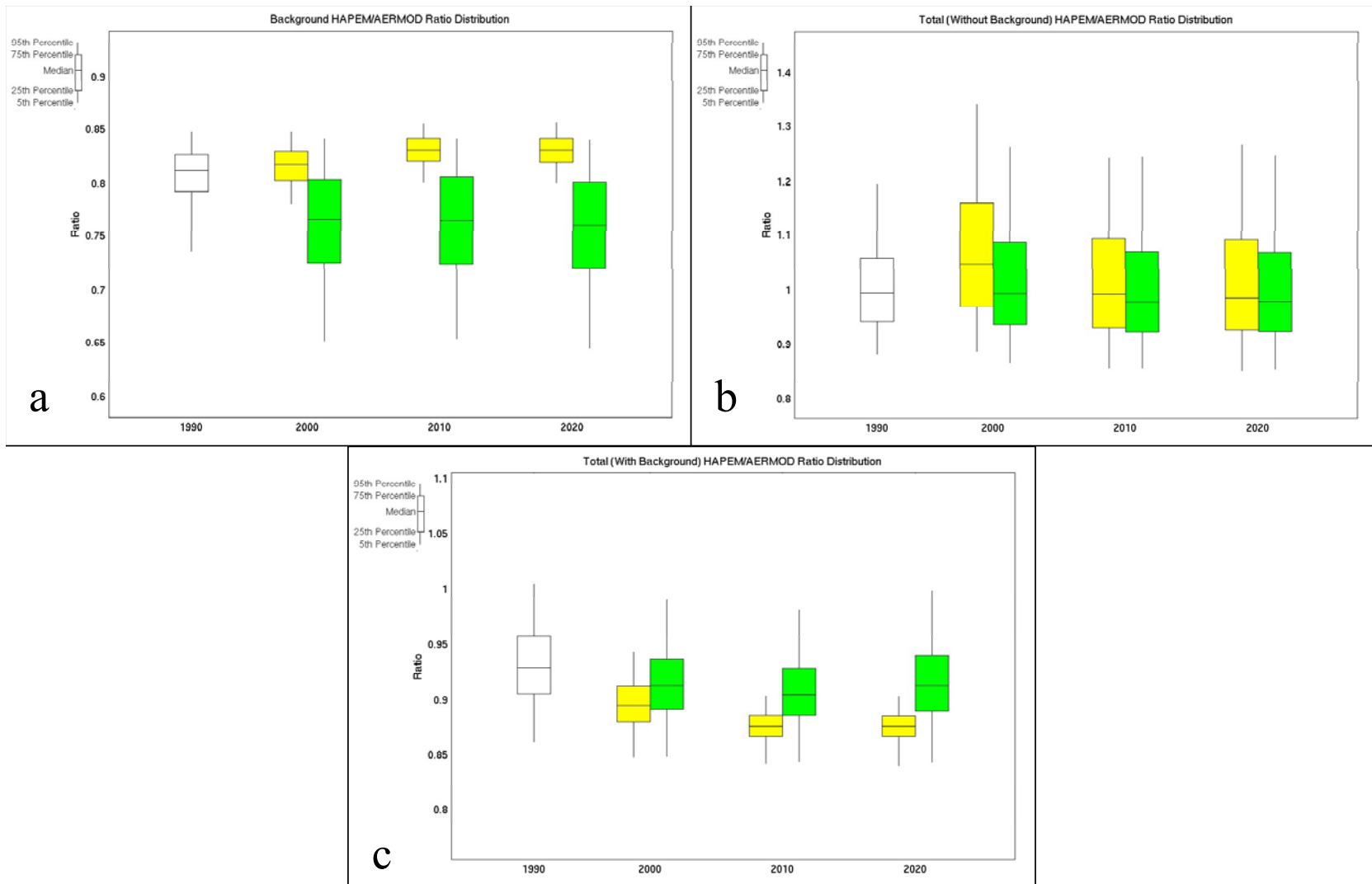
Figure 42. Box and whisker plots of tract level HAPEM6/AERMOD ratio distributions for 1990, 2000, 2010, and 2020 for a) background b) all sources not including background, and c) all sources. Box colors are the same as in Figure 29.

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 406XXXXXXX 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 407XXXXXXX 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

B-5

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

B-7

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

B-9

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		

INDUSTRIAL ECONOMICS, INCORPORATED

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. (See the summary table in Attachment 1 for specific strengths and weaknesses of the individual studies).

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,

---

[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/).

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).

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 x 10^{-6} to 7.8 x 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.

<table>
<tr><th colspan="3" align="center">EXHIBIT 3
COMPARISON OF THE PLIOFILM COHORT AND THE CHINESE WORKER COHORT</th></tr>
<tr><td></td><td>Pliofilm Cohort</td><td>Chinese Worker Cohort</td></tr>
<tr><td>Description of Industry</td><td>Workers in Pliofilm manufacturing plants in two locations in Ohio</td><td>Workers in 672 factories in 12 cities of China employed in a number of industries such as painting, printing, footwear, rubber, and chemical</td></tr>
<tr><td>Cohort Size/Number of Leukemia Cases</td><td>1,717 white males/14 cases of leukemia</td><td>74,828 benzene exposed workers/47 cases of leukemia</td></tr>
<tr><td>Dates of Employment/ Follow-up</td><td>1939-1976/Follow-up through 1987 Rinsky et al. (2002) followed subjects through 1996.</td><td>1972-1987</td></tr>
<tr><td>Exposure Levels with Positive Effects</td><td>>40 ppm (cumulative exposure)</td><td><10 ppm (average exposure); <40 ppm-years (cumulative exposure)</td></tr>
<tr><td>Exposure Assessment Method</td><td>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.</td><td>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).</td></tr>
<tr><td>Major Results</td><td>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).</td><td>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).</td></tr>
<tr><td>Strengths</td><td>-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</td><td>-Larger number of cases of leukemia
-Positive results seen at lower benzene exposures</td></tr>
<tr><td>Limitations</td><td>-Relatively smaller number of cases of leukemia
-Benzene exposures higher than those experienced by the general public
-No measurement data available prior to 1946.</td><td>-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</td></tr>
</table>

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. *[Placeholder: Refinement of the cessation lag structure may also have implications for our selection of a risk model to use in the health benefits analysis. Ideally, we would choose a model with an exposure window and lag most consistent with the proposed cessation lag. We will update this section to address this issue as the cessation lag structure evolves.]*

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) <u>Epidemiology Kept Simple: An Introduction to Classic and Modern Epidemiology.</u> 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.

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.

IEc

FIGURE D-1:   LIFE-TABLE MODEL OVERVIEW
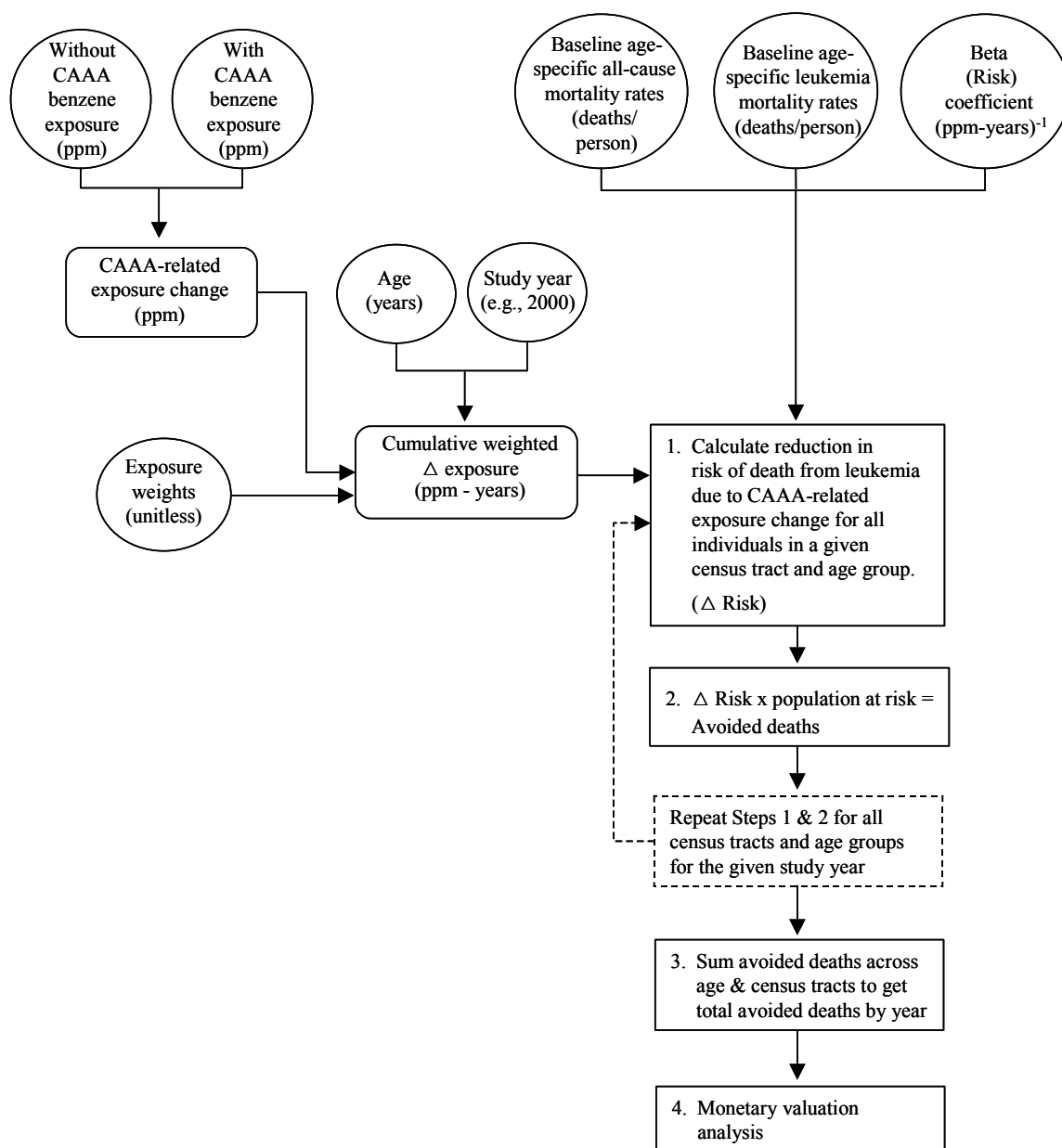
Note: This flowchart assumes the model is being run with leukemia mortality data. The model can also be run with leukemia incidence data. The difference between the model results for these two runs represents an estimate of avoided non-fatal cases of leukemia.

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: $$\Delta E_{i,j,k} = \Sigma \left(\Delta e_{i,j,k} \times w_t \right)$$
$$j = \max(j\text{-}i^*,1995) \text{ to } j; \ i = 0 \text{ to } i; \text{ and } t = 0 \text{ to } j - \max(j\text{-}i,1995)$$

* 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: $\qquad\qquad R^e_{i,j,k} - R^o_i = \Delta R_{i,j,k}$

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

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:

$$R_g = (E_{g\ Without\text{-}CAAA} - E_{g\ With\text{-}CAAA})/E_{g\ Without\text{-}CAAA}$$

Where:

R_g = percent reduction in emissions occurring within attached garages due to the CAAA;

$E_{g\ Without\text{-}CAAA}$ = total emissions occurring within attached garages under the *Without-CAAA* scenario in tons/year ($E_{g\ non\text{-}road\ Without} + E_{g\ on\text{-}road\ Without}$); and

$E_{g\ With\text{-}CAAA}$ = total emissions occurring within attached garages under the *With-CAAA* scenario in tons/year ($E_{g\ non\text{-}road\ With} + E_{g\ on\text{-}road\ With}$).

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.	

<u>On-road Emissions Occurring Within Attached Garages</u>

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 five minutes 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 estimates 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.